

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 45 Seconds

(without alignments)  
2475.065 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353

Sequence: 1 MTGPGHKCECKSHYVDGL.....ALAAYSFRINRKTIGXFHF 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size: 9

Total number of hits satisfying chosen parameters: 12

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database:

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	27.8	1192	4 Q9H7H7	Q9h7h7 homo sapien
2	98	27.8	1416	4 Q86UR4	Q86ur4 homo sapien
3	98	27.8	1736	4 Q8TES1	Q8tes1 homo sapien
4	98	27.8	2551	4 Q8WQ08	Q8wq08 homo sapien
5	98	27.8	2551	4 Q8IUG9	Q8iug9 homo sapien
6	98	27.8	2551	4 Q7Z5N9	Q7z5n9 homo sapien
7	93	26.3	897	4 Q9NRV3	Q9nrv3 homo sapien
8	93	26.3	1069	4 Q9UF98	Q9uf98 homo sapien
9	32	9.1	500	6 Q864U4	Q864u4 bos taurus
10	30	8.5	894	11 Q8BM87	Q8bm87 mus musculu
11	30	8.5	2559	11 Q8R4U0	Q8r4u0 mus musculu
12	25	7.1	1431	11 Q8CFM6	Q8cfm6 rattus norv

#### ALIGNMENTS

RESULT 1  
Q9H7H7 PRELIMINARY; PRT; 1192 AA.  
ID Q9H7H7  
AC Q9H7H7  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE FLJ00112 protein (Fragment).  
GN FLJ00112.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=spleen;  
RA Ohara O., Nagase T., Kikuno R., Okumura K.;  
RT "The nucleotide sequence of a long cDNA clone isolated from human spleen."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024503; BAB15793.1; -  
DR HSP; P98066; ITSG.  
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR InterPro; IPR000782; BIGH3 FAS1.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR000538; Link.  
DR Pfam; PF00008; EGF; 7.  
DR Pfam; PF02469; Fasciclin; 2.  
DR Pfam; PF00193; Xlink; 1.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR PRINTS; PR01265; LINKMODULE.  
DR PRODOM; PD000918; Link; 1.  
DR SMART; SM00181; EGF; 9.  
DR SMART; SM00554; FAS1; 3.  
DR SMART; SM00445; LINK; 1.  
DR PROSITE; PS00022; EGF\_1; 4.  
DR PROSITE; PS01186; EGF\_2; 8.  
DR PROSITE; PS02113; FAS1; 3.  
DR PROSITE; PS01248; LAMININ TYPE\_EGF; 2.  
KW EGF-like domain; Laminin EGF-like domain.  
FT NON TER 1  
SQ SEQUENCE 1192 AA; 128738 MW; 5966BED4B83BD9C1 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 1192;  
Best Local Similarity 100.0%; Pred. No. 1.4e-94;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVMSFPSLTNFLTFLVAYNSSSARGAFLEHLTDLISIRGTLFVPONSGLGENET 228  
Db 952 SGNLLQVMSFPSLTNFLTFLVAYNSSSARGAFLEHLTDLISIRGTLFVPONSGLGENET 1011  
Qy 229 LSGRDIHHLANVSMFFYNDLVNGTTLQTRLGSKLIT 266  
Db 1012 LSGRDIHHLANVSMFFYNDLVNGTTLQTRLGSKLIT 1049

#### RESULT 2

Q86UR4 PRELIMINARY; PRT; 1416 AA.  
ID Q86UR4  
AC Q86UR4  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hyaluronan receptor for endocytosis (Fragment).  
GN HARE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

EMBL; AK074051; BAB84877.1; -  
GO; GO:0005540; F:hyaluronic acid binding; IEA.  
GO; GO:0005198; F:structural molecule activity; IEA.  
GO; GO:0007155; P:cell adhesion; IEA.  
InterPro; IPR000782; BIGH3\_FAS1.  
InterPro; IPR006209; EGF\_like.  
InterPro; IPR006210; IEGF.  
InterPro; IPR002049; Laminin\_EGF.  
Pfam; PF00008; EGF; 12.  
Pfam; PF02469; Fasciclin; 4.  
Pfam; PF00193; Xlink; 1.  
PRINTS; PR00011; EGFLAMININ.  
ProDom; PD000918; Link; 1.  
SMART; SM00181; EGF; 15.  
SMART; SM00554; FAS1; 5.  
SMART; SM00445; LINK; 1.  
PROSITE; PS00022; EGF\_1; 4.  
PROSITE; PS01196; EGF\_2; 12.  
PROSITE; PS0213; FAST; 5.  
PROSITE; PS01248; LAMININ\_TYPE\_EGF; 2.  
EGF-like domain.  
NON\_TER  
FT  
SEQUENCE 1736 AA; 187887 MW; 50982047E43925F2 CRC64;  
3

Query Match 27.8%; Score 98; DB 4; Length 1736;  
Best Local Similarity 100.0%; Pred.No. 1.9e-94;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SGNLLQLVMSFFSLNLFTEVLAYSNSARGRAFLHLEHTDLSIRGTLFPQNSGLGENET 238  
Db 1496 SGNLLQLVMSFFSLNLFTEVLAYSNSARGRAFLHLEHTDLSIRGTLFPQNSGLGENET 1555

QY 229 LSGRDIEHHLANVSFFYNLDVNGTTLQTRLGSKLLIT 266  
Db 1556 LSGRDIEHHLANVSFFYNLDVNGTTLQTRLGSKLLIT 1593

RESULT 4

ID Q8WQ08 PRELIMINARY; PRT; 2551 AA.  
QB Q8WQ08;  
AC Q8WQ08;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DD 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DS 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
ST Stablein-2.  
AB STAB2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
POLYT O.; Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,  
JOHANSSON S., Birk R., Hakky N., Franke P., Kodella V., Kamnits C.,  
ORFANOS C.B., JOHANSSON S., Goerdts S.;  
"Stabilin-1 and stabilin-2 constitute a novel family of fasciclin  
domain-containing adhesion molecules associated with endothelial-  
macrophage differentiation and angiogenic processes.";  
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
EMBL; AJ295695; CACB2105.1; -  
GO; GO:0005540; F:hyaluronic acid binding; IEA.  
GO; GO:0005198; F:structural molecule activity; IEA.  
GO; GO:0007155; P:cell adhesion; IEA.  
InterPro; IPR000782; BIGH3\_FAS1.  
InterPro; IPR006209; EGF\_like.  
InterPro; IPR006210; IEGF.  
InterPro; IPR002049; Laminin\_EGF.  
InterPro; IPR00538; Link.  
Pfam; PF00008; EGF; 16.  
Pfam; PF02469; Fasciclin; 5.  
Pfam; PF00193; Xlink; 1.

```
DR PRINTS; PRO00011; EGF, LAMININ.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 20.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS02013; FAS1; 7.
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
KW EGF-like domain; Laminin EGF-like domain.
SQ SEQUENCE 2551 AA; 276992 MW; 60A44651CC2B69 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 2551;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 2311 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 2370

Qy 229 LSGRDIEHHLANVMFFYNDLVNGTTTLQTRLGSKLLIT 266
Db 2371 LSGRDIEHHLANVMFFYNDLVNGTTTLQTRLGSKLLIT 2408

RESULT 5
Q8IU99 PRELIMINARY; PRT; 2551 AA.
AC Q8IU99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FELE-2.
GN FELE-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Adachi H., Tsujimoto M.;
RT "FELE-1, a novel scavenger receptor with in vitro bacteria-binding and
RT angiogenesis-modulating activities."
RL J. Biol. Chem. 277:34264-34270 (2002).
DR EMBL; AB052958; BAC15608.1; -;
DR GenBank; U000000000; STAB2.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR001881; EGF CA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; LEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00008; EGF; 16.
DR Pfam; PF02469; Fasciclin; 6.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PRO0011; EGF, LAMININ.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 22.
DR SMART; SM00179; EGF CA; 8.
DR SMART; SM00180; EGF_Lam; 5.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR SMART; SM00286; FTL; 8.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS02013; FAS1; 7.
```

```
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
DR SQ SEQUENCE 2551 AA; 277026 MW; 3ADD3F5BD2019C4 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 2551;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 2311 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 2370

Qy 229 LSGRDIEHHLANVMFFYNDLVNGTTTLQTRLGSKLLIT 266
Db 2371 LSGRDIEHHLANVMFFYNDLVNGTTTLQTRLGSKLLIT 2408

RESULT 6
Q7Z5N9 PRELIMINARY; PRT; 2551 AA.
AC Q7Z5N9;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FEX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Park S.-Y., Kim I.-S.;
RT "FEX2, a novel cell adhesion molecule of Fas-1 superfamily mediates
RT cell-cell interaction."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY311388; AAF74958.1; -;
DR SQ SEQUENCE 2551 AA; 276986 MW; 3ACB6A6C3CB80044 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 2551;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 2311 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 2370

Qy 229 LSGRDIEHHLANVMFFYNDLVNGTTTLQTRLGSKLLIT 266
Db 2371 LSGRDIEHHLANVMFFYNDLVNGTTTLQTRLGSKLLIT 2408

RESULT 7
Q9NR33 PRELIMINARY; PRT; 897 AA.
AC Q9NR33;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD44-like precursor FELL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tao Q., Zhang W., Cao X.;
RT "Molecular cloning and characterization of human FELL sharing homology
RT with CD44."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF160476; AAF82398.1; -;
DR HSP; P98066; ITSG.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

DR InterPro: IPR000782; BIGH3\_FAS1.  
DR InterPro: IPR006209; EGF-like.  
DR InterPro: IPR006210; IEGF.  
DR InterPro: IPR002049; Laminin\_EGF.  
DR InterPro: IPR000538; Link.  
DR Pfam: PF00008; EGF; 4.  
DR Pfam: PF02469; Fasciclin; 2.  
DR Pfam: PF00193; Xlink; 1.  
DR PRINTS: PR01265; LINKMODULE.  
DR ProDom: PD000918; Link; 1.  
DR SMART: SM00181; EGF; 5.  
DR SMART: SM00554; FAS1; 3.  
DR SMART: SM00445; LINK; 1.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 4.  
DR PROSITE: PS02113; FAS1; 3.  
DR PROSITE: PS01248; LAMININ\_TYPE\_EGF; 1.  
KW EGF-like domain.  
SQ SEQUENCE 897 AA; 97585 MW; EB920AF36101E388 CRC64;  
Query Match 26.3%; Score 93; DB 4; Length 897;  
Best Local Similarity 100.0%; Pred. No. 2.2e-89;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 60  
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 548  
Qy 61 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 93  
Db 549 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 581  
RESULT 8  
Q9UF98 PRELIMINARY; PRT; 1069 AA.  
ID Q9UF98  
AC Q9UF98  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE Hypothetical protein (fragment).  
GN DKF2P434E0321.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL133021; CAB61358.1; -;  
DR F1R; T42681; T42681.  
DR HSSP; P98056; I7SG.  
DR GO; GO:0005540; F-hyaluronic acid binding; IEA.  
DR GO; GO:0005198; F-structural molecule activity; IEA.  
DR GO; GO:0007155; P-cell adhesion; IEA.  
DR InterPro: IPR000782; BIGH3\_FAS1.  
DR InterPro: IPR006209; EGF-like.  
DR InterPro: IPR006210; IEGF.  
DR InterPro: IPR002049; Laminin\_EGF.  
DR Pfam: PF00008; EGF; 7.  
DR Pfam: PF02469; Fasciclin; 2.  
DR PRINTS: PR00193; Xlink; 1.  
DR PRINTS: PR00011; EGF\_LAMININ.  
DR PRINTS: PR01265; LINKMODULE.  
DR ProDom: PD000918; Link; 1.  
DR SMART: SM00181; EGF; 8.  
DR SMART: SM00554; FAS1; 3.  
DR SMART: SM00445; LINK; 1.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 5.

DR PROSITE: PS02113; FAS1; 3.  
DR PROSITE: PS01248; LAMININ\_TYPE\_EGF; 1.  
KW Hypothetical protein; EGF-like domain; Laminin EGF-like domain.  
FT NON\_TER 1  
SQ SEQUENCE 1069 AA; 115666 MW; 04B0960932164D63 CRC64;  
Query Match 26.3%; Score 93; DB 4; Length 1069;  
Best Local Similarity 100.0%; Pred. No. 2.5e-99;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 60  
Db 697 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 756  
Qy 61 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 93  
Db 757 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 789  
RESULT 9  
Q864U4 PRELIMINARY; PRT; 500 AA.  
ID Q864U4  
AC Q864U4  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative hyaluronan receptor for endocytosis (fragment).  
GN HARE.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Ulbrich S.E., Thoenes S.;  
RL "Expression of the hyaluronan system in the bovine oviduct.";  
RN Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Einspanner R.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ550060; CAD79334.2; -;  
KW Receptor.  
FT NON\_TER 1  
FT NON\_TER 500  
SQ SEQUENCE 500 AA; 54479 MW; C93AA6404B103582 CRC64;  
Query Match 9.1%; Score 32; DB 6; Length 500;  
Best Local Similarity 100.0%; Pred. No. 8.2e-25;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 48 LHFQDTTGVGFHLRSLPGQYKLTDFDKAREACA 79  
Db 422 LHFQDTTGVGFHLRSLPGQYKLTDFDKAREACA 453  
RESULT 10  
Q8BM87 PRELIMINARY; PRT; 894 AA.  
ID Q8BM87  
AC Q8BM87  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to CD44-like precursor FELL (fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.



```

RC STRAIN=C57BL/6J; TISSUE=Dienecephalon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60, 770 full-length cDNAs";
RL Nature 420:563-573(2002);
DR EMBL; AK034522; BAC28741.1; -.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF02469; Fasciclin; 2.
DR PRINTS; PR01265; LINKMODULE.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00180; EGF Lam; 2.
DR SMART; SM00554; FAS1; 3.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS0213; FAS1; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
FT NON TER 1
SQ SEQUENCE 894 AA; 97059 MW; 328652A223CFBB96 CRC64;

Query Match 8.5%; Score 30; DB 11; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.9e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 74
Db 530 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 559

RESULT 11
Q8R4U0 PRELIMINARY; PRT; 2559 AA.
AC Q8R4U0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Stabilin-2.
GN STAB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=21818465; PubMed=11829752;
RA Politz O., Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,
RA Johansson S., Svineng G., Franke P., Kannicht C., Kzyshkowska J.,
RA Longati P., Velten F.W., Johansson S., Goerdts S.,
RT "Stabilin-1 and -2 constitute a novel family of fasciclin-like
RT hyaluronan receptor homologues";
RL Biochem. J. 362:155-164(2002).
DR EMBL; AF364951; AAL91684.2; -.
DR MGD; MGI:2178743; Stab2.
DR GO; GO:003779; F:actin binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001589; Actbind actnin.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR006209; EGF-like.

```

```

DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 13.
DR Pfam; PF02469; Fasciclin; 6.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 21.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00019; ACTININ 1; 1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS0213; FAS1; 7.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW EGF-like domain.
SQ SEQUENCE 2559 AA; 277530 MW; 1C9855AD61EFF015 CRC64;

Query Match 8.5%; Score 30; DB 11; Length 2559;
Best Local Similarity 100.0%; Pred. No. 5e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 74
Db 2195 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 2224

RESULT 12
Q8CFM6 PRELIMINARY; PRT; 1431 AA.
AC Q8CFM6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hyaluronan receptor for endocytosis HARE precursor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22169209; PubMed=12181351;
RA Zhou B., Weigel J.A., Saxena A., Weigel P.H.;
RT "Molecular Cloning and Functional Expression of the Rat 175-kDa
RT Hyaluronan Receptor for Endocytosis.";
RL Mol. Biol. Cell 13:2853-2868(2002).
DR EMBL; AY007370; AAG13634.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF02469; Fasciclin; 3.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00179; EGF CA; 2.
DR SMART; SM00554; FAS1; 4.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS0213; FAS1; 4.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.

```

KW Receptor.  
 FT NON\_TER 1  
 SQ SEQUENCE 1431 AA; 155914 MW; 3340A8F7AF366722 CRC64;  
 Query Match 7.1%; Score 25; DB 11; Length 1431;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-17;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 50 FQDTTVGVFHLRSPGLGQYKLTFDKA 74  
 DB 1057 FQDTTVGVFHLRSPGLGQYKLTFDKA 1081

Search completed: April 1, 2004, 09:58:24  
 Job time : 46 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 17 Seconds  
(without alignments)  
1081.221 Million cell updates/sec

Title: US-09-466-778B-11  
Perfect score: 353  
Sequence: 1 MTGPGKHCKCKSHYVGDGL.....ALAAYSYRINRKTIGFXHF 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size: 9

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

No matches found

Search completed: April 1, 2004, 09:57:25  
Job time : 17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 20 Seconds  
(without alignments)  
1697.780 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353

Sequence: 1 MTGPGKHCKECKSHYVGDL.....ALAAYSYFRINRKTIGEXHF 353

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size: 9

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database: PIR.78.\*

1: Pirl.\*

2: Pirl2.\*

3: Pirl3.\*

4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	26.3	1069	2 T42681	hypothetical prote

#### ALIGNMENTS

#### RESULT 1

T42681

hypothetical protein DKFP434E0321.1 - human (fragments)

C/Species: Homo sapiens (man)

C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000

C/Accession: T42681

R/Blum, H.; Baurerachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, November 1999

A/Reference number: Z22233

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-870;871-1069 <AAA>

A/Cross-references: EMBL:AL13021

A/Experimental source: adult testis; clone DKFP434E0321

A/Note: the cDNA sequence contains a +1 frameshift near codon 870

C/Genetics:

A/Note: DKFP434E0321.1

Query Match

Best Local Similarity 100.0%; Pred.No. 6.6e-88; Length 1069;

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECKSHYVGDLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDTTGVGPHL 60  
Db 597 MTGPGKHCKECKSHYVGDLNCEPEQLPDRCLQDNGOCHADAKCVDLHFQDTTGVGPHL 756

Qy 61 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 93

Db 757 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 789

Search completed: April 1, 2004, 09:58:56  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:58:28 ; Search time 45 Seconds  
(without alignments)  
2057.758 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353

Sequence: 1 MTGPKHKCKSHYVDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1069545 seqs, 262320428 residues

Word size : 9

Total number of hits satisfying chosen parameters: 23

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	27.8	1192	15	US-10-028-248A-40
2	98	27.8	1192	15	US-10-107-782-40
3	98	27.8	1394	9	US-09-842-930A-25
4	98	27.8	1416	14	US-10-133-172-4
5	98	27.8	1653	14	US-10-133-172-20
6	93	26.3	897	15	US-10-028-248A-41
7	93	26.3	897	15	US-10-028-248A-211
8	93	26.3	897	15	US-10-107-782-41
9	93	26.3	897	15	US-10-107-782-211
10	93	26.3	1069	15	US-10-028-248A-39
11	93	26.3	1069	15	US-10-107-782-39
12	93	26.3	2420	15	US-10-028-248A-4
13	93	26.3	2420	15	US-10-107-782-4
14	93	26.3	2675	15	US-10-028-248A-2
15	93	26.3	2675	15	US-10-107-782-2

16 54 15.3 93 14 US-10-133-172-5  
17 25 7.1 1431 9 US-09-842-930A-2  
18 25 7.1 1431 14 US-10-133-172-2  
19 20 5.7 106 14 US-10-023-896-82  
20 18 5.1 18 14 US-10-133-172-17  
21 15 4.2 15 9 US-09-842-930A-55  
22 13 3.7 13 9 US-09-842-930A-43  
23 13 3.7 13 14 US-10-133-172-18

## ALIGNMENTS

RESULT 1  
US-10-028-248A-40  
; Sequence 40, Application US/10028248A  
; Publication No. US20030235882A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Vernet, Corine  
; APPLICANT: Casman, Stacie  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Miller, Charles  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Colman, Steven  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Si, Jingsheng  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Stone, David  
; APPLICANT: Sciore, Paul  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Rothenberg, Mark  
; TITLE OF INVENTION: No. US20030235882A1elel Nucleic Acids and Polypeptides and Methods  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 21402-222  
; CURRENT APPLICATION NUMBER: US/10/028,248A  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/256619  
; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 60/262959  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/272408  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/285189  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 60/308039  
; PRIOR FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: 60/311266  
; PRIOR FILING DATE: 2001-08-09  
; NUMBER OF SEQ ID NOS: 211  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 40  
; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-028-248A-40

Query Match 27.8%; Score 98; DB 15; Length 1192;  
Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SGNLLVMSFSLNLTFLTEVLAYSNSARGAFLEHLDLSIRGTLFVPONSGLNET 228

Db 952 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1011  
 Qy 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266  
 Db 1012 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1049

RESULT 2

US-10-107-782-40  
 ; Sequence 40, Application US/10107782  
 ; Publication No. US20040018970A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boldog, Ferenc  
 ; APPLICANT: Casman, Stacie  
 ; APPLICANT: Colman, Steve  
 ; APPLICANT: Edinger, Shlomit  
 ; APPLICANT: Gangolli, Esna  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Liu, Xiaohong  
 ; APPLICANT: Malyankar, Uriel  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Rothenberg, Mark  
 ; APPLICANT: Sciore, Paul  
 ; APPLICANT: Shenoy, Suresh  
 ; APPLICANT: Shinkets, Richard  
 ; APPLICANT: Si, Jingsheng  
 ; APPLICANT: Smithson, Glenda  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Stone, David  
 ; APPLICANT: Taupier, Raymond, Jr.  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Zerhusen, Brian  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 21402-222CIP  
 ; CURRENT APPLICATION NUMBER: US/10/107,782  
 ; CURRENT FILING DATE: 2002-03-27  
 ; PRIOR APPLICATION NUMBER: 10/028,248  
 ; PRIOR FILING DATE: 2001-12-19  
 ; PRIOR APPLICATION NUMBER: 60/256,619  
 ; PRIOR FILING DATE: 2000-12-19  
 ; PRIOR APPLICATION NUMBER: 60/262,959  
 ; PRIOR FILING DATE: 2001-01-19  
 ; PRIOR APPLICATION NUMBER: 60/272,408  
 ; PRIOR FILING DATE: 2001-02-28  
 ; PRIOR APPLICATION NUMBER: 60/285,189  
 ; PRIOR FILING DATE: 2001-04-20  
 ; PRIOR APPLICATION NUMBER: 60/308,039  
 ; PRIOR FILING DATE: 2001-07-26  
 ; PRIOR APPLICATION NUMBER: 60/311,266  
 ; PRIOR FILING DATE: 2001-08-09  
 ; PRIOR APPLICATION NUMBER: 60/279,344  
 ; PRIOR FILING DATE: 2001-03-28  
 ; NUMBER OF SEQ ID NOS: 215  
 ; SOFTWARE: Curaseq1 version 0.1  
 ; SEQ ID NO 40  
 ; LENGTH: 1192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-10-107-782-40

Query Match 27.8%; Score 98; DB 15; Length 1192;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228  
 Db 952 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1011  
 Qy 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266

Db 1012 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1049

RESULT 3

US-09-842-930A-25  
 ; Sequence 25, Application US/09842930A  
 ; Publication No. US20020197681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Weigel, Paul  
 ; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis  
 ; FILE REFERENCE: 5820.603  
 ; CURRENT APPLICATION NUMBER: US/09/842,930A  
 ; CURRENT FILING DATE: 2001-04-22  
 ; PRIOR APPLICATION NUMBER: 60/245,320  
 ; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: 60/199,538  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 56  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 25  
 ; LENGTH: 1394  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-842-930A-25  
 Query Match 27.8%; Score 98; DB 9; Length 1394;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-89;  
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228  
 Db 1154 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1213  
 Qy 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266  
 Db 1214 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1251

RESULT 4

US-10-133-172-4  
 ; Sequence 4, Application US/10133172  
 ; Publication No. US20030104987A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WEIGEL, PAUL H  
 ; APPLICANT: WEIGEL, JANET A  
 ; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS  
 ; FILE REFERENCE: 5864.014  
 ; CURRENT APPLICATION NUMBER: US/10/133,172  
 ; CURRENT FILING DATE: 2002-04-25  
 ; PRIOR APPLICATION NUMBER: 60/286,468  
 ; PRIOR FILING DATE: 2001-04-25  
 ; PRIOR APPLICATION NUMBER: 09/842,930  
 ; PRIOR FILING DATE: 2001-04-25  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 1416  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-133-172-4  
 Query Match 27.8%; Score 98; DB 14; Length 1416;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-89;  
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228  
 Db 1176 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1235  
 Qy 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266  
 Db 1236 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1273

```

; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-41

Query Match          26.3%;   Score 93;   DB 15;   Length 897;
Best Local Similarity 100.0%;   Pred. No. 1.5e-84;
Matches 93;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0

QY      1  MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
Db      489  MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 548

QY      61  RSPLGQYKLTFFDKAREACANEAAATMATYNQLSY 93
Db      549  RSPLGQYKLTFFDKAREACANEAAATMATYNQLSY 581

RESULT 7
US-10-028-248A-211
; Sequence 211, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189

```

;; PRIOR FILING DATE: 2001-04-20  
;; PRIOR APPLICATION NUMBER: 60/308039  
;; PRIOR FILING DATE: 2001-07-26  
;; PRIOR APPLICATION NUMBER: 60/311266  
;; PRIOR FILING DATE: 2001-08-09  
;; NUMBER OF SEQ ID NOS: 211  
;; SOFTWARE: Patentin Ver. 2.1  
;; SEQ ID NO 211  
;; LENGTH: 897  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-028-248A-211

Query Match 26.3%; Score 93; DB 15; Length 897;  
Best Local Similarity 100.0%; Pred. No. 1.5e-84;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60  
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 548  
QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSY 93  
Db 549 RSPGQYKLTDFDKAREACANEATMATYNQLSY 581

RESULT 8  
US-10-107-782-41  
;; Sequence 41, Application US/10107782  
;; Publication No. US20040018970A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Boldog, Ferenc,  
;; APPLICANT: Casman, Stacie,  
;; APPLICANT: Colman, Steve,  
;; APPLICANT: Edinger, Shlomit,  
;; APPLICANT: Gangolli, Esha,  
;; APPLICANT: Kekuda, Ramesh,  
;; APPLICANT: Li, Li,  
;; APPLICANT: Liu, Xiaohong,  
;; APPLICANT: Malyankar, Uriel,  
;; APPLICANT: Miller, Charles,  
;; APPLICANT: Millet, Isabelle,  
;; APPLICANT: Patturajan, Meera,  
;; APPLICANT: Rothenberg, Mark,  
;; APPLICANT: Sciore, Paul,  
;; APPLICANT: Shenoy, Suresh,  
;; APPLICANT: Shimkets, Richard,  
;; APPLICANT: Si, Jingsheng,  
;; APPLICANT: Smithson, Glennda,  
;; APPLICANT: Spytek, Kimberly,  
;; APPLICANT: Stone, David,  
;; APPLICANT: Taupier, Raymond, Jr.,  
;; APPLICANT: Tchernev, Velizar,  
;; APPLICANT: Vernet, Corine,  
;; APPLICANT: Zerhusen, Brian

;; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF  
;; FILE REFERENCE: 21402-222CIP  
;; CURRENT APPLICATION NUMBER: US/10/107,782  
;; CURRENT FILING DATE: 2002-03-27  
;; PRIOR APPLICATION NUMBER: 10/028,248  
;; PRIOR FILING DATE: 2001-12-19  
;; PRIOR APPLICATION NUMBER: 60/256,619  
;; PRIOR FILING DATE: 2000-12-19  
;; PRIOR APPLICATION NUMBER: 60/262,959  
;; PRIOR FILING DATE: 2001-01-19  
;; PRIOR APPLICATION NUMBER: 60/272,408  
;; PRIOR FILING DATE: 2001-02-28  
;; PRIOR APPLICATION NUMBER: 60/285,189  
;; PRIOR FILING DATE: 2001-04-20  
;; PRIOR APPLICATION NUMBER: 60/308,039  
;; PRIOR FILING DATE: 2001-07-26  
;; PRIOR APPLICATION NUMBER: 60/311,266  
;; PRIOR FILING DATE: 2001-08-09

;; PRIOR APPLICATION NUMBER: 60/279,344  
;; PRIOR FILING DATE: 2001-03-28  
;; NUMBER OF SEQ ID NOS: 215  
;; SOFTWARE: Curaseqlist version 0.1  
;; SEQ ID NO 41  
;; LENGTH: 897  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-107-782-41

Query Match 26.3%; Score 93; DB 15; Length 897;  
Best Local Similarity 100.0%; Pred. No. 1.5e-84;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60  
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 548  
QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSY 93  
Db 549 RSPGQYKLTDFDKAREACANEATMATYNQLSY 581

RESULT 9  
US-10-107-782-211  
;; Sequence 211, Application US/10107782  
;; Publication No. US20040018970A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Boldog, Ferenc,  
;; APPLICANT: Casman, Stacie,  
;; APPLICANT: Colman, Steve,  
;; APPLICANT: Edinger, Shlomit,  
;; APPLICANT: Gangolli, Esha,  
;; APPLICANT: Kekuda, Ramesh,  
;; APPLICANT: Li, Li,  
;; APPLICANT: Liu, Xiaohong,  
;; APPLICANT: Malyankar, Uriel,  
;; APPLICANT: Miller, Charles,  
;; APPLICANT: Millet, Isabelle,  
;; APPLICANT: Patturajan, Meera,  
;; APPLICANT: Rothenberg, Mark,  
;; APPLICANT: Sciore, Paul,  
;; APPLICANT: Shenoy, Suresh,  
;; APPLICANT: Shimkets, Richard,  
;; APPLICANT: Si, Jingsheng,  
;; APPLICANT: Smithson, Glennda,  
;; APPLICANT: Spytek, Kimberly,  
;; APPLICANT: Stone, David,  
;; APPLICANT: Taupier, Raymond, Jr.,  
;; APPLICANT: Tchernev, Velizar,  
;; APPLICANT: Vernet, Corine,  
;; APPLICANT: Zerhusen, Brian

;; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF  
;; FILE REFERENCE: 21402-222CIP  
;; CURRENT APPLICATION NUMBER: US/10/107,782  
;; CURRENT FILING DATE: 2002-03-27  
;; PRIOR APPLICATION NUMBER: 10/028,248  
;; PRIOR FILING DATE: 2001-12-19  
;; PRIOR APPLICATION NUMBER: 60/256,619  
;; PRIOR FILING DATE: 2000-12-19  
;; PRIOR APPLICATION NUMBER: 60/262,959  
;; PRIOR FILING DATE: 2001-01-19  
;; PRIOR APPLICATION NUMBER: 60/272,408  
;; PRIOR FILING DATE: 2001-02-28  
;; PRIOR APPLICATION NUMBER: 60/285,189  
;; PRIOR FILING DATE: 2001-04-20  
;; PRIOR APPLICATION NUMBER: 60/308,039  
;; PRIOR FILING DATE: 2001-07-26  
;; PRIOR APPLICATION NUMBER: 60/311,266  
;; PRIOR FILING DATE: 2001-08-09  
;; PRIOR APPLICATION NUMBER: 60/279,344  
;; PRIOR FILING DATE: 2001-03-28  
;; NUMBER OF SEQ ID NOS: 215



```
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 211
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-211

Query Match      26.3%; Score 93; DB 15; Length 897;
Best Local Similarity 100.0%; Pred. No. 1.5e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
Db 489 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 548

QY 61 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 93
Db 549 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 581

RESULT 10
US-10-028-248A-39
; Sequence 39, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uziel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glenda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Coleman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-39
```

```
Query Match      26.3%; Score 93; DB 15; Length 1069;
Best Local Similarity 100.0%; Pred. No. 1.8e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
Db 697 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 756

QY 61 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 93
Db 757 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 789

RESULT 11
US-10-107-782-39
; Sequence 39, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc
; APPLICANT: Casman, Stacie
; APPLICANT: Coleman, Steve
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gangolli, Esha
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uziel
; APPLICANT: Miller, Charles
; APPLICANT: Millet, Isabelle
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Sciore, Paul
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard
; APPLICANT: Si, Jingsheng
; APPLICANT: Smithson, Glenda
; APPLICANT: Spytek, Kimberly
; APPLICANT: Stone, David
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US/10/107,782
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 39
; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-39
```

```
Query Match      26.3%; Score 93; DB 15; Length 1069;
Best Local Similarity 100.0%; Pred. No. 1.8e-84;
```

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60  
 Db 697 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 756  
 QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSY 93  
 Db 757 RSPGQYKLTDFDKAREACANEATMATYNQLSY 789

RESULT 12  
 US-10-028-248A-4  
 ; Sequence 4, Application US/10028248A  
 ; Publication No. US20030235882A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shinkets, Richard  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Casman, Stacie  
 ; APPLICANT: Malyankar, Uriel  
 ; APPLICANT: Shenoy, Suresh  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Gangolli, Esha  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Boldog, Ferenc  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Taupier Jr, Raymond J  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Zernhusen, Bryan  
 ; APPLICANT: Liu, Xiaohong  
 ; APPLICANT: Colman, Steven  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Si, Jingsheng  
 ; APPLICANT: Edinger Shlomit  
 ; APPLICANT: Stone, David  
 ; APPLICANT: Sciore, Paul  
 ; APPLICANT: Miller, Isabelle  
 ; APPLICANT: Rothenberg, Mark  
 ; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods  
 ; FILE REFERENCE: 21402-222  
 ; CURRENT APPLICATION NUMBER: US/10/028,248A  
 ; PRIOR FILING DATE: 2001-12-19  
 ; PRIOR APPLICATION NUMBER: 60/256619  
 ; PRIOR FILING DATE: 2000-12-19  
 ; PRIOR APPLICATION NUMBER: 60/262959  
 ; PRIOR FILING DATE: 2001-01-19  
 ; PRIOR APPLICATION NUMBER: 60/272408  
 ; PRIOR FILING DATE: 2001-02-28  
 ; PRIOR APPLICATION NUMBER: 60/285189  
 ; PRIOR FILING DATE: 2001-04-20  
 ; PRIOR APPLICATION NUMBER: 60/308039  
 ; PRIOR FILING DATE: 2001-07-26  
 ; PRIOR APPLICATION NUMBER: 60/311266  
 ; PRIOR FILING DATE: 2001-08-09  
 ; NUMBER OF SEQ ID NOS: 211  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 2420  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (371)..(372)  
 ; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid  
 US-10-028-248A-4

Query Match 26.3%; Score 93; DB 15; Length 2420;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-84;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60  
 Db 2250 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 2309  
 QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSY 93  
 Db 2310 RSPGQYKLTDFDKAREACANEATMATYNQLSY 2342

RESULT 13  
 US-10-107-782-4  
 ; Sequence 4, Application US/10107782  
 ; Publication No. US20040018970A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boldog, Ferenc  
 ; APPLICANT: Casman, Stacie  
 ; APPLICANT: Colman, Steve  
 ; APPLICANT: Edinger, Shlomit  
 ; APPLICANT: Gangolli, Esha  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Liu, Xiaohong  
 ; APPLICANT: Malyankar, Uriel  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Rothenberg, Mark  
 ; APPLICANT: Sciore, Paul  
 ; APPLICANT: Shenoy, Suresh  
 ; APPLICANT: Shinkets, Richard  
 ; APPLICANT: Si, Jingsheng  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Stone, David  
 ; APPLICANT: Taupier, Raymond, jr.,  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Zernhusen, Brian  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 21402-222CIP  
 ; CURRENT APPLICATION NUMBER: US/10/107,782  
 ; CURRENT FILING DATE: 2002-03-27  
 ; PRIOR APPLICATION NUMBER: 10/028,248  
 ; PRIOR FILING DATE: 2001-12-19  
 ; PRIOR APPLICATION NUMBER: 60/256,619  
 ; PRIOR FILING DATE: 2000-12-19  
 ; PRIOR APPLICATION NUMBER: 60/262,959  
 ; PRIOR FILING DATE: 2001-01-19  
 ; PRIOR APPLICATION NUMBER: 60/272,408  
 ; PRIOR FILING DATE: 2001-02-28  
 ; PRIOR APPLICATION NUMBER: 60/285,189  
 ; PRIOR FILING DATE: 2001-04-20  
 ; PRIOR APPLICATION NUMBER: 60/308,039  
 ; PRIOR FILING DATE: 2001-07-26  
 ; PRIOR APPLICATION NUMBER: 60/311,266  
 ; PRIOR FILING DATE: 2001-08-09  
 ; PRIOR APPLICATION NUMBER: 60/279,344  
 ; PRIOR FILING DATE: 2001-03-28  
 ; NUMBER OF SEQ ID NOS: 215  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 4  
 ; LENGTH: 2420  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (371)..(371)  
 ; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (372)..(372)  
 ; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid  
 US-10-107-782-4

```

Query Match      26.3%; Score 93; DB 15; Length 2420;
Best Local Similarity 100.0%; Pred. No. 3.8e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
    |||||
Db 2250 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 2309

QY 61 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 93
    |||||
Db 2310 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 2342

```

```

RESULT 14
US-10-028-248A-2
; Sequence 2, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Miller, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2675
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (55)..(56)
; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (427)..(428)

```

```

; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid
US-10-028-248A-2

Query Match      26.3%; Score 93; DB 15; Length 2675;
Best Local Similarity 100.0%; Pred. No. 4.1e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
    |||||
Db 2303 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 2362

QY 61 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 93
    |||||
Db 2363 RSPLGQYKLTDFDKAREACANEAAATMATYNQLSY 2395

```

```

RESULT 15
US-10-107-782-2
; Sequence 2, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gangolli, Esha
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel
; APPLICANT: Miller, Charles
; APPLICANT: Millet, Isabelle
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Sciore, Paul
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard
; APPLICANT: Si, Jingsheng
; APPLICANT: Smithson, Glennnda
; APPLICANT: Spytek, Kimberly
; APPLICANT: Stone, David
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 2
; LENGTH: 2675
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```
; LOCATION: (55)..(55)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (56)..(56)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (427)..(427)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-107-782-2

Query Match          26.3%; Score 93; DB 15; Length 2675;
Best Local Similarity 100.0%; Pred. No. 4.1e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60
Db 2303 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 2362

Qy 61 RSPLGQKLTDFDKAREACANEATMATVYNQLSY 93
Db 2363 RSPLGQKLTDFDKAREACANEATMATVYNQLSY 2395

RESULT 16
US-10-133-172-5
; Sequence 5, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT FILING DATE: 2002-04-25
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-5

Query Match          15.3%; Score 54; DB 14; Length 93;
Best Local Similarity 100.0%; Pred. No. 3.8e-46;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 QKAKYHLCAGLWETGRVAYPTAFASQCGSVGVIVDYGPRPNKSEMWDFCY 148
Db 40 QKAKYHLCAGLWETGRVAYPTAFASQCGSVGVIVDYGPRPNKSEMWDFCY 93

RESULT 17
US-09-842-930A-2
; Sequence 2, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.503
; CURRENT FILING DATE: 2001-04-22
; PRIOR FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
```

```
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-842-930A-2

Query Match          7.1%; Score 25; DB 9; Length 1431;
Best Local Similarity 100.0%; Pred. No. 7.6e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 FQDTTVGVFHLRSPGLGQYKLTDFDKA 74
Db 1057 FQDTTVGVFHLRSPGLGQYKLTDFDKA 1081

RESULT 18
US-10-133-172-2
; Sequence 2, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT FILING DATE: 2002-04-25
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-133-172-2

Query Match          7.1%; Score 25; DB 14; Length 1431;
Best Local Similarity 100.0%; Pred. No. 7.6e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 FQDTTVGVFHLRSPGLGQYKLTDFDKA 74
Db 1057 FQDTTVGVFHLRSPGLGQYKLTDFDKA 1081

RESULT 19
US-10-023-896-82
; Sequence 82, Application US/10023896
; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004p1
; CURRENT FILING DATE: 2001-12-21
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/023,896
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 106
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-896-82

Query Match      5.7%; Score 20; DB 14; Length 106;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 IILVTGAVALAAYSYFNR 345
Db 23 IILVTGAVALAAYSYFNR 42

RESULT 20
US-10-133-172-17
; Sequence 17, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; TITLE OF INVENTION: IDENTIFICATION OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-17

Query Match      5.1%; Score 18; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPGKHCKCKSHVVDG 19
Db 1 TPGKHCKCKSHVVDG 18

RESULT 21
US-09-842-930A-55
; Sequence 55, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; PRIOR FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-930A-55

Query Match      4.2%; Score 15; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 RSPLGQYKLTFFDKAR 75
Db 1 RSPLGQYKLTFFDKAR 15

RESULT 22
US-09-842-930A-43
; Sequence 43, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; PRIOR FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-930A-43

Query Match      3.7%; Score 13; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 GTLFVPQNSGLGE 225
Db 1 GTLFVPQNSGLGE 13

RESULT 23
US-10-133-172-18
; Sequence 18, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-18

Query Match      3.7%; Score 13; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 PIDRCLQDNGQCH 40
Db 1 PIDRCLQDNGQCH 13

Search completed: April 1, 2004, 10:04:13
Job time : 46 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 22 Seconds  
(without alignments)  
828.362 Million cell updates/sec

Title: US-09-466-778B-11  
Perfect score: 353  
Sequence: 1 MTGPGKHCKCKSHYVGDL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size: 9

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued Patents AA.\*  
1: /cgm2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgm2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
---------------	-------	----------------	--------	-------	-------------

No matches found

Search completed: April 1, 2004, 09:59:31  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:42 ; Search time 58 Seconds  
(without alignments)  
1719.643 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353

Sequence: 1 MTGPGKHCKECKSHYVDGL.....ALAAYSYFRINRKTIGXFHF 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size: 9

Total number of hits satisfying chosen parameters: 20

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : A Geneseq\_29Jan04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	345	97.7	333	3	AAY93913
2	98	27.8	1394	5	AAM47684 Human Hya
3	98	27.8	1416	6	ABG72499 Human 190
4	98	27.8	1653	6	ABG72514 Human 190
5	93	26.3	244	4	AAB83359 NOV3 prot
6	93	26.3	669	4	AAB83364 NOV8 prot
7	93	26.3	669	4	AAB83362 NOV6 prot
8	93	26.3	897	5	ABJ10588 Human nov
9	93	26.3	2420	5	ABJ10587 Human nov
10	93	26.3	2675	5	ABJ10586 Human nov
11	85	24.1	330	3	AAB42164 Human ORF
12	57	16.1	315	4	AAB83358 NOV2 prot
13	54	15.3	33	6	ABG72500 Human 190
14	25	7.1	1431	5	AAM47675 Rat Hyalu
15	25	7.1	1431	6	ABG72498 Rat 175kd
16	20	5.7	106	4	AUU00893 Human can
17	18	5.1	18	6	ABG72512 Human 190
18	15	4.2	15	5	AAM47702 Human HAR
19	13	3.7	13	5	AAM47690 Human HAR
20	13	3.7	13	6	ABG72513 Human 190

# ALIGNMENTS

RESULT 1  
AAY93913  
ID AAY93913 standard; protein; 353 AA.  
XX  
AC AAY93913;  
XX  
DT 03-OCT-2000 (first entry)  
XX  
DE A human hyaluronan-binding protein, designated BM-HABP.  
XX  
KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;  
proliferative condition; metastasis; inflammation; ischemia;  
host defence dysfunction; immune surveillance dysfunction; arthritis;  
multiple sclerosis; autoimmunity; immune dysfunction; allergy.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Misc-difference 94 /note= "unspecified amino acid encoded by NNC"  
FT Domain 121..215 /note= "HA binding domain"  
FT Misc-difference 157 /note= "unspecified amino acid encoded by TNC"  
FT Misc-difference 303 /note= "unspecified amino acid encoded by YCT"  
FT Misc-difference 314 /note= "unspecified amino acid encoded by NCC"  
FT Misc-difference 320 /note= "unspecified amino acid encoded by GNA"  
FT Misc-difference 324 /note= "unspecified amino acid encoded by TNT"  
FT Misc-difference 325 /note= "unspecified amino acid encoded by GNC"  
FT Misc-difference 351 /note= "unspecified amino acid encoded by CAN"  
XX  
WO200039166-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 20-DEC-1999; 99WO-US030462.  
XX  
PR 23-DEC-1998; 98US-0113871P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (AMNA-) AMERICAN NAT RED CROSS.  
XX  
FI Hastings GA, Liau G, Tsifrina E;  
XX  
WPI; 2000-452376/39.  
DR N-PSDB; AAA57365.  
XX  
FT New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-HABP and BM-HABP, useful for treating proliferative conditions, metastasis, inflammation, ischemia, arthritis and multiple sclerosis.  
PS Claim 11; Fig 4A-B; 457pp; English.  
XX  
CC The present sequence represents a hyaluronan-binding protein. The specification describes four hyaluronan-binding proteins, known as WF-HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for treating diseases such as proliferative conditions, metastasis, inflammation, ischemia, host defence dysfunction, immune surveillance dysfunction, arthritis, multiple sclerosis, autoimmunity, immune dysfunction and allergy  
XX  
SQ Sequence 353 AA;  
Query Match 97.7%; Score 345; DB 3; Length 353;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKESHYVGDGLNCPPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60  
Db 1 MTGPGKHCKESHYVGDGLNCPPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60  
QY 61 RSLPGQKLTDFDKAREACANEATMATYNOLSYXQAKYHLCSAGMLETGRVAYPTAFAS 120  
Db 61 RSLPGQKLTDFDKAREACANEATMATYNOLSYXQAKYHLCSAGMLETGRVAYPTAFAS 120  
QY 121 QNCGSGVVGIVDYGPRNPKSEMDVFCYRMKDVNCTKVGVDGSGSYSNLLQVLMSPF 180  
Db 121 QNCGSGVVGIVDYGPRNPKSEMDVFCYRMKDVNCTKVGVDGSGSYSNLLQVLMSPF 180  
QY 181 SLTNFLTEVLAYSNSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHLAN 240  
Db 181 SLTNFLTEVLAYSNSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHLAN 240  
QY 241 VSMFFYNDLVNGTTLQTRLSKLLITDRQDPLHPTETRCVDRDGTLEWDICASNGITHVI 300  
Db 241 VSMFFYNDLVNGTTLQTRLSKLLITDRQDPLHPTETRCVDRDGTLEWDICASNGITHVI 300  
QY 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353  
Db 301 SRXLKAPPAPVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353

RESULT 2  
AAM47684  
ID AAM47684 standard; protein; 1394 AA.  
XX AAM47684;  
AC  
DT 22-FEB-2002 (first entry)  
XX  
DE Human Hyaluronic Acid Receptor for Endocytosis, HARE.  
XX  
KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;  
KW chondroitin sulphate; extracellular matrix; cartilage; skin;  
KW vitreous humour; endocytic receptor; glycosaminoglycan; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200181544-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 25-APR-2001; 2001WO-US013403.  
XX  
PR 25-APR-2000; 2000US-0199538P.  
PR 02-NOV-2000; 2000US-0245320P.  
XX  
XX (WEIG/) WEIGEL P A.  
PA (ZHOU/) ZHOU B.  
PA (WEIG/) WEIGEL J A.  
XX  
XX Weigel PA, Zhou B, Weigel JA;  
XX  
XX WPI; 2002-049271/06.  
XX  
XX New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for  
PT identifying agents that inhibit binding to hyaluronic acid, and related  
PT nucleic acid.  
XX  
XX Claim 20; Fig 33; 263pp; English.  
XX  
XX The present invention relates to sequences for rat and human HARE  
CC (Hyaluronic Acid Receptor for Endocytosis, ABA04548, ABA04562, AAM47675  
CC and AAM47684). HARE can bind specifically to at least one of hyaluronic  
CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin  
CC sulphate (CS). HA is an extracellular matrix component of all tissues, in  
CC particular cartilage, skin and vitreous humour. HARE is the endocytic  
CC receptor responsible for removing HA and other glycosaminoglycans from  
CC the circulation. The present sequence is human HARE

SQ Sequence 1394 AA;  
Query Match 27.8%; Score 98; DB 5; Length 1394;  
Best Local Similarity 100.0%; Pred. No. 3.4e-87;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENET 228  
Db 1154 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENET 1213  
QY 229 LSGRDIIEHLANVSFFYNDLVNGTTLQTRLSKLLIT 266  
Db 1214 LSGRDIIEHLANVSFFYNDLVNGTTLQTRLSKLLIT 1251  
RESULT 3  
ABG72499  
ID ABG72499 standard; protein; 1416 AA.  
XX  
AC ABG72499;  
XX  
DT 19-FEB-2003 (first entry)  
XX  
DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #1.  
XX  
KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
KW tumour; gene therapy; human; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200286093-A2.  
XX  
PD 31-OCT-2002.  
XX  
XX 25-APR-2002; 2002WO-US013209.  
XX  
PR 25-APR-2001; 2001US-00842930.  
PR 25-APR-2001; 2001US-0286468P.  
XX  
XX (WEIG/) WEIGEL P H.  
PA (WEIG/) WEIGEL J A.  
XX  
XX Weigel PH, Weigel JA;  
XX  
XX WPI; 2003-093126/08.  
XX  
XX N-PSDB; ABX13822.  
XX  
XX Targeting compounds e.g. chemotherapeutic agent to cell of subject  
PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE,  
PT epitope.  
XX  
XX Example; Fig 9A; 167pp; English.  
XX  
XX The invention describes a method of targeting a compound to a cell or  
CC tissue of an individual expressing a functionally active hyaluronan (HA)  
CC receptor for endocytosis (HARE) or a cell that does not express  
CC functionally active HARE. The method involves using HA molecule, a  
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
CC method is useful for targeting a compound, preferably a chemotherapeutic  
CC agent or a radioisotope to cell of an individual, especially a human,  
CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
CC method useful for preventing interaction between a cell having at least  
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
CC its surface. This second method is useful for preventing metastasis by  
CC preventing interaction between tumour cells having HA, CD or CDS coat and  
CC non-tumour cells expressing HARE on its surface. The invention also  
CC describes a method useful for detecting the presence of HA, CD and CDS in  
CC a biological fluid. This sequence encodes the human 190kDa Hyaluronan  
CC receptor for endocytosis (HARE)



SQ Sequence 1416 AA;  
Query Match 27.8%; Score 98; DB 6; Length 1416;  
Best Local Similarity 100.0%; Pred. No. 3.4e-87;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228  
DB 1176 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1235  
QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 266  
DB 1236 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 1273

RESULT 4  
ABG72514  
ID ABG72514 standard; protein; 1653 AA.  
XX AC ABG72514;  
XX AC ABG72514;  
DT 19-FEB-2003 (first entry)  
DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #2.  
XX KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
KW tumour; gene therapy; human.  
XX OS Homo sapiens.  
XX PN WO200286093-A2.  
XX PD 31-OCT-2002.  
XX PF 25-APR-2002; 2002WO-US013209.  
XX PR 25-APR-2001; 2001US-00842930.  
XX PR 25-APR-2001; 2001US-0286468P.  
XX PA (WEIG/) WEIGEL P H.  
XX PA (WEIG/) WEIGEL J A.  
XX PI Weigel PH, Weigel JA;  
XX DR WPI; 2003-093126/08.  
XX DR N-PSDB; ABX13823.  
XX PT Targeting compounds e.g. chemotherapeutic agent to cell of subject  
PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
PT epitope.  
XX PS Example; Fig 9B; 167pp; English.

XX CC The invention describes a method of targeting a compound to a cell or  
CC tissue of an individual expressing a functionally active hyaluronan (HA)  
CC receptor for endocytosis (HARE) or a cell that does not express  
CC functionally active HARE. The method involves using HA molecule, a  
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
CC method is useful for targeting a compound, preferably a chemotherapeutic  
CC agent or a radioisotope to cell of an individual, especially a human,  
CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
CC method useful for preventing interaction between a cell having at least  
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
CC its surface. This second method is useful for preventing metastasis by  
CC preventing interaction between tumour cells having HA, CD or CDS coat and  
CC non-tumour cells expressing HARE on its surface. The invention also  
CC describes a method useful for detecting the presence of HA, CD and CDS in  
CC a biological fluid. This is the amino acid sequence of a longer version  
CC of the human 190kDa Hyaluronan receptor for endocytosis (HARE) shown in  
CC ABG72499 (Encoded by ABX13822)

XX SQ Sequence 1653 AA;  
Query Match 27.8%; Score 98; DB 6; Length 1653;  
Best Local Similarity 100.0%; Pred. No. 3.9e-87;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228  
DB 1413 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1472  
QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 266  
DB 1473 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 1510

RESULT 5  
AAB83359  
ID AAB83359 standard; protein; 244 AA.  
XX AC AAB83359;  
XX AC AAB83359;  
DT 26-MAR-2002 (first entry)  
DE NOV3 protein sequence.  
XX KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;  
KW cardiovascular; casein kinase II phosphorylation site; contraception;  
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;  
KW epidermal growth factor; cell development; apoptosis; cell adhesion;  
KW growth migration; cell structure; motility; cancer; immune disorder;  
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;  
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;  
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;  
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.  
XX OS Unidentified.  
XX PN WO200136638-A2.  
XX PD 25-MAY-2001.  
XX PF 17-NOV-2000; 2000WO-US031543.  
XX PR 19-NOV-1999; 99US-0166336P.  
XX PR 29-NOV-1999; 99US-0167785P.  
XX PR 08-MAR-2000; 2000US-0187844P.  
XX PR 16-NOV-2000; 2000US-00715417.  
XX PA (CURA-) CURAGEN CORP.  
XX PI Shimkets RA, Lichenstein H, Vernet C, Fernandes E;  
XX DR WPI; 2001-648134/74.  
XX DR N-PSDB; AAF87114.  
XX PT Novel human polypeptides and the nucleic acids that encode them useful  
XX for preventing, diagnosing and treating e.g. cancer, inflammation and  
XX immune disorders.  
XX PS Claim 1; Page 14-17; 141pp; English.

XX CC This sequence is the NOV3 protein. The invention relates to the NOV1-  
CC NOV16 proteins, and their coding sequences. The proteins have Cytostatic;  
CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular  
CC activities. The sequences may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate NOV3 expression. They  
CC may be used to treat disorders associated with decreased expression by  
CC rectifying mutations or deletions in a patient's genome that affect the  
CC activity of protein by expressing inactive proteins or to supplement the  
CC patients own production of protein. They are used to produce NOV3  
CC proteins, by inserting the nucleic acid into a cell and culturing it to  
CC express the protein. The DNA may be used as DNA probes in assays to

CC detect and quantitate the presence of similar DNAs in samples, and which  
 CC patients may need restorative therapy. The NOVX protein may also be used  
 CC as antigens in the production of antibodies (Abs) against NOVX and in  
 CC assays to identify modulators of NOVX expression and activity. The anti-  
 CC NOVX Abs and antagonist are used to down regulate expression and  
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX  
 CC in samples. Disorders that may be prevented, diagnosed and/or treated  
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and  
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of  
 CC serine/threonine kinases, and are used to treat kinase-related disorders  
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contractions).  
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)  
 CC -like super family and are involved in, e.g. regulation of cell  
 CC development, apoptosis, cell adhesion, growth migration, cell structure  
 CC and motility and protein management, and are used to treat cancers,  
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.  
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to  
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT  
 CC syndrome and marfan syndrome  
 CC  
 CC Sequence 244 AA;

Query Match 26.3%; Score 93; DB 4; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-83;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60  
 Db 145 MTGPGKHCKCKSHYVGDGLNCEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 204  
 QY 61 RSPGLQYKLTDFKAREACANEATMATYNQLSY 93  
 Db 205 RSPGLQYKLTDFKAREACANEATMATYNQLSY 237

RESULT 6  
 AAB83364  
 ID AAB83364 standard; protein; 669 AA.  
 AC AAB83364;  
 XX  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX NOV8 protein sequence.

NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;  
 cardiovascular; casein kinase II phosphorylation site; contraception;  
 serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;  
 epidermal growth factor; cell development; apoptosis; cell adhesion;  
 growth migration; cell structure; motility; cancer; immune disorder;  
 inflammatory disorder; cellular adhesion disorder; long-QT syndrome;  
 cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;  
 therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;  
 NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.

Unidentified.  
 WO200136638-A2.  
 XX  
 XX 25-MAY-2001.  
 XX  
 XX 17-NOV-2000; 2000WO-US031543.  
 XX  
 XX 19-NOV-1999; 99US-0166336P.  
 XX 29-NOV-1999; 99US-0167785P.  
 XX 08-MAR-2000; 2000US-0187844P.  
 XX 16-NOV-2000; 2000US-00715417.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Lichenstein H, Vernet C, Fernandes E;  
 XX WPI; 2001-648134/74.  
 DR

DR N-PSDB; AAF871119.  
 XX  
 PT Novel human polypeptides and the nucleic acids that encode them useful  
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and  
 PT immune disorders.  
 XX  
 XX  
 PS Claim 1; Page 29-30; 141pp; English.

CC This sequence is the NOV8 protein. The invention relates to the NOV1-  
 CC NOV16 proteins, and their coding sequences. The proteins have Cytostatic;  
 CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular  
 CC activities. The sequences may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate NOVX expression. They  
 CC may be used to treat disorders associated with decreased expression by  
 CC rectifying mutations or deletions in a patient's genome that affect the  
 CC activity of protein by expressing inactive proteins or to supplement the  
 CC patients own production of protein. They are used to produce NOVX  
 CC proteins, by inserting the nucleic acid into a cell and culturing it to  
 CC express the protein. The DNA may be used as DNA probes in assays to  
 CC detect and quantitate the presence of similar DNAs in samples, and which  
 CC patients may need restorative therapy. The NOVX protein may also be used  
 CC as antigens in the production of antibodies (Abs) against NOVX and in  
 CC assays to identify modulators of NOVX expression and activity. The anti-  
 CC NOVX Abs and antagonist are used to down regulate expression and  
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX  
 CC in samples. Disorders that may be prevented, diagnosed and/or treated  
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and  
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of  
 CC serine/threonine kinases, and are used to treat kinase-related disorders  
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraction).  
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)  
 CC -like super family and are involved in, e.g. regulation of cell  
 CC development, apoptosis, cell adhesion, growth migration, cell structure  
 CC and motility and protein management, and are used to treat cancers,  
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.  
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to  
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT  
 CC syndrome and marfan syndrome  
 CC  
 CC Sequence 669 AA;

Query Match 26.3%; Score 93; DB 4; Length 669;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-82;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60  
 Db 566 MTGPGKHCKCKSHYVGDGLNCEPEQLPDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 625  
 QY 61 RSPGLQYKLTDFKAREACANEATMATYNQLSY 93  
 Db 626 RSPGLQYKLTDFKAREACANEATMATYNQLSY 658

RESULT 7  
 AAB83362  
 ID AAB83362 standard; protein; 669 AA.  
 XX  
 XX AAB83362;  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX NOV6 protein sequence.

NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;  
 cardiovascular; casein kinase II phosphorylation site; contraception;  
 serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;  
 epidermal growth factor; cell development; apoptosis; cell adhesion;  
 growth migration; cell structure; motility; cancer; immune disorder;  
 inflammatory disorder; cellular adhesion disorder; long-QT syndrome;  
 cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;  
 therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;  
 NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.

XX OS Unidentified.  
 XX PN WO200136638-A2.  
 XX PD 25-MAY-2001.  
 XX PF 17-NOV-2000; 2000WO-US031543.  
 XX PR 19-NOV-1999; 99US-0166336P.  
 XX PR 29-NOV-1999; 99US-0167789P.  
 XX PR 08-MAR-2000; 2000US-0187844P.  
 XX PR 16-NOV-2000; 2000US-00715417.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Lichenstein H, Vernet C, Fernandes E;  
 XX DR WPI; 2001-648134/74.  
 XX DR N-PSDB; AAF87117.  
 XX FT Novel human polypeptides and the nucleic acids that encode them useful  
 XX FT for preventing, diagnosing and treating e.g. cancer, inflammation and  
 XX FT immune disorders.  
 XX PS Claim 1; Page 24-25; 141pp; English.  
 XX CC This sequence is the NOV6 protein. The invention relates to the NOV1-  
 XX CC NOV16 proteins, and their coding sequences. The proteins have cytostatic;  
 XX CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular  
 XX CC activities. The sequences may be used in the prevention, diagnosis and  
 XX CC treatment of diseases associated with inappropriate NOVX expression. They  
 XX CC may be used to treat disorders associated with decreased expression by  
 XX CC rectifying mutations or deletions in a patient's genome that affect the  
 XX CC activity of protein by expressing inactive proteins or to supplement the  
 XX CC patients own production of protein. They are used to produce NOVX  
 XX CC proteins, by inserting the nucleic acid into a cell and culturing it to  
 XX CC express the protein. The DNA may be used as DNA probes in assays to  
 XX CC detect and quantitate the presence of similar DNAs in samples, and which  
 XX CC patients may need restorative therapy. The NOVX protein may also be used  
 XX CC as antigens in the production of antibodies (Abs) against NOVX and in  
 XX CC assays to identify modulators of NOVX expression and activity. The anti-  
 XX CC NOVX Abs and antagonist are used to down regulate expression and  
 XX CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX  
 XX CC in samples. Disorders that may be prevented, diagnosed and/or treated  
 XX CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and  
 XX CC NOV13-16 have casein kinase II phosphorylation sites characteristic of  
 XX CC serine/threonine kinases, and are used to treat kinase-related disorders  
 XX CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).  
 XX CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)  
 XX CC -like super family and are involved in, e.g. regulation of cell  
 XX CC development, apoptosis, cell adhesion, growth migration, cell structure  
 XX CC and motility and protein management, and are used to treat cancers,  
 XX CC inflammatory disorders, immune disorders and cellular adhesion disorders.  
 XX CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to  
 XX CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT  
 XX CC syndrome and marfan syndrome  
 XX SQ Sequence 669 AA;  
 Query March 26.3%; Score 93; DB 4; Length 669;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-02;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 60  
 Db 566 MTGPGKHCKECSHYVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 625  
 Qy 61 RSPGLQYKLTDFKAREACANEAATMATYNQLSY 93  
 Db 626 RSPGLQYKLTDFKAREACANEAATMATYNQLSY 658

RESULT 8  
 ABJ10588  
 ID ABJ10588 standard; protein; 897 AA.  
 XX AC ABJ10588;  
 XX DT 28-NOV-2002 (first entry)  
 XX DE Human novel protein NOV1c SEQ ID NO: 211.  
 XX KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;  
 KW antiarteriosclerotic; antidiabetic; antiaschmatic; antiinflammatory;  
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;  
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;  
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;  
 KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;  
 KW antiaddictive.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 221 /note= "optionally Cys depending on polymorphism present  
 FT in coding sequence"  
 FT Misc-difference 325 /note= "optionally Leu depending on polymorphism present  
 FT in coding sequence"  
 FT Misc-difference 416 /note= "optionally Tyr depending on polymorphism present  
 FT in coding sequence"  
 FT Misc-difference 832 /note= "optionally Gly depending on polymorphism present  
 FT in coding sequence"  
 XX EN WO200259315-A2.  
 XX PD 01-AUG-2002.  
 XX PF 19-DEC-2001; 2001WO-US050076.  
 XX PR 19-DEC-2000; 2000US-0256619P.  
 XX PR 19-JAN-2001; 2001US-0262959P.  
 XX PR 28-FEB-2001; 2001US-0272408P.  
 XX PR 20-APR-2001; 2001US-0285189P.  
 XX PR 26-JUL-2001; 2001US-0308039P.  
 XX PR 09-AUG-2001; 2001US-0311266P.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyaukar U;  
 PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;  
 PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;  
 PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;  
 PI Rothenberg M;  
 XX DR WPI; 2002-666903/71.  
 XX DR N-PSDB; ABT08489.  
 XX PT New isolated NOVX polypeptides and polynucleotides, useful for  
 XX FT preventing, diagnosing or treating NOVX-associated disorders e.g.  
 XX FT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease  
 XX FT or Alzheimer's disease.  
 XX PS Claim 54; Page 25; 363pp; English.  
 XX CC The present invention provides the protein and coding sequences of  
 XX CC several novel human proteins, designated NOVX. These can be used in the  
 XX CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-  
 XX CC Lindau syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,  
 XX CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral  
 XX CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia  
 XX CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,  
 XX CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,

CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,  
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,  
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or  
 CC graft-versus-host disease. The present sequence is a protein of the  
 CC invention  
 XX  
 SQ Sequence 897 AA;  
 Query Match 26.3%; Score 93; DB 5; Length 897;  
 Best Local Similarity 100.0%; Pred. No. 2e-82;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60  
 Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 548  
 QY 61 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 93  
 Db 549 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 581  
 RESULT 9  
 ABJ10587  
 ID ABJ10587 standard; protein; 2420 AA.  
 XX  
 AC ABJ10587;  
 XX  
 DT 28-NOV-2002 (first entry)  
 XX  
 DE Human novel protein NOV1b SEQ ID NO: 4.  
 XX  
 KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;  
 KW antiarteriosclerotic; antidiabetic; antiaslathmic; antiinflammatory;  
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;  
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;  
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;  
 KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;  
 KW antiaddictive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200259315-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 19-DEC-2001; 2001WO-US050076.  
 XX  
 PR 19-DEC-2000; 2000US-0256619P.  
 XX  
 PR 19-JAN-2001; 2001US-0262959P.  
 XX  
 PR 28-FEB-2001; 2001US-0272408P.  
 XX  
 PR 20-APR-2001; 2001US-0285189P.  
 XX  
 PR 26-JUL-2001; 2001US-0308039P.  
 XX  
 PR 09-AUG-2001; 2001US-0311266P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;  
 PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;  
 PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;  
 PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;  
 PI Rothenberg M;  
 XX  
 DR WPI; 2002-666903/71.  
 DR N-PSDB; ABT08488.  
 XX  
 PT New isolated NOVX polypeptides and polynucleotides, useful for  
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.  
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease  
 PT or Alzheimer's disease.  
 XX  
 PS Claim 1; Page 23-24; 363pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of

CC several novel human proteins, designated NOVX. These can be used in the  
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-  
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,  
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral  
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia  
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,  
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,  
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,  
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,  
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or  
 CC graft-versus-host disease. The present sequence is a protein of the  
 CC invention  
 XX  
 SQ Sequence 2420 AA;  
 Query Match 26.3%; Score 93; DB 5; Length 2420;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-82;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60  
 Db 2250 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 2309  
 QY 61 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 93  
 Db 2310 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 2342  
 RESULT 10  
 ABJ10586  
 ID ABJ10586 standard; protein; 2675 AA.  
 XX  
 AC ABJ10586;  
 XX  
 DT 28-NOV-2002 (first entry)  
 XX  
 DE Human novel protein NOV1a SEQ ID NO: 2.  
 XX  
 KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;  
 KW antiarteriosclerotic; antidiabetic; antiaslathmic; antiinflammatory;  
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;  
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;  
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;  
 KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;  
 KW antiaddictive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200259315-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 19-DEC-2001; 2001WO-US050076.  
 XX  
 PR 19-DEC-2000; 2000US-0256619P.  
 XX  
 PR 19-JAN-2001; 2001US-0262959P.  
 XX  
 PR 28-FEB-2001; 2001US-0272408P.  
 XX  
 PR 20-APR-2001; 2001US-0285189P.  
 XX  
 PR 26-JUL-2001; 2001US-0308039P.  
 XX  
 PR 09-AUG-2001; 2001US-0311266P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;  
 PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;  
 PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;  
 PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;  
 PI Rothenberg M;  
 XX  
 DR WPI; 2002-666903/71.  
 DR N-PSDB; ABT08487.  
 XX  
 PT New isolated NOVX polypeptides and polynucleotides, useful for

PT preventing, diagnosing or treating NOVX-associated disorders e.g.  
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease  
XX or Alzheimer's disease.

PS Claim 1; Page 20; 363pp; English.

XX The present invention provides the protein and coding sequences of  
CC several novel human proteins, designated NOVX. These can be used in the  
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-  
CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,  
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral  
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia  
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,  
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,  
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,  
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis, or  
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or  
CC graft-versus-host disease. The present sequence is a protein of the  
CC invention  
XX

XX Sequence 2675 AA;

Query Match 26.3%; Score 93; DB 5; Length 2675;  
Best Local Similarity 100.0%; Pred. No. 4.9e-82;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIRCLQDNGQCHADAKCVLHFDQTTVGVFHL 60

Db 2303 MTGPGKHCKCKSHYVGDGLNCEPEQLPIRCLQDNGQCHADAKCVLHFDQTTVGVFHL 2362

QY 61 RSPILGQYKLTDFDKAREACANEATMATYNQLSY 93

Db 2363 RSPILGQYKLTDFDKAREACANEATMATYNQLSY 2395

RESULT 11

ID AAB42164  
ID AAB42164 standard; protein; 330 AA.

XX AAB42164;

XX 08-FEB-2001 (first entry)

XX Human ORFX ORF1928 polypeptide sequence SEQ ID NO:3856.

XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;  
KW vulnery; antipsoriatic; antiparkinsonian; nontropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antithyroid; antihypertensive;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

PA (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC76373.

XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 3007-3008; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
CC antipsoriatic; antiparkinsonian; nontropic; neuroprotective; osteopathic;  
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antithyroid; antihypertensive; antianemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX

SQ Sequence 330 AA;

Query Match 24.1%; Score 85; DB 3; Length 330;

Best Local Similarity 100.0%; Pred. No. 6.4e-75;

Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SGNLLQVLMSPSLTNFLTEVLAYSNSRARGFLAHLTDLSTRTGLFVFPQNSGLGENET 228

Db 89 SGNLLQVLMSPSLTNFLTEVLAYSNSRARGFLAHLTDLSTRTGLFVFPQNSGLGENET 148

QY 229 LSGRDIEHLANVSMFFYNDLVNGT 253

Db 149 LSGRDIEHLANVSMFFYNDLVNGT 173

RESULT 12

AAB83358

ID AAB83358 standard; protein; 315 AA.

XX AAB83358;

XX 26-MAR-2002 (first entry)

XX NOV2 protein sequence.

XX NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;  
KW cardiovascular; casein kinase II phosphorylation site; contraception;  
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;  
KW epidermal growth factor; cell development; apoptosis; cell adhesion;  
KW growth migration; cell structure; motility; cancer; immune disorder;  
KW inflammatory disorder; cellular adhesion disorder; long-OT syndrome;  
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;  
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;  
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.

XX Unidentified.

XX WO200136638-A2.

PD 25-MAY-2001.  
 XX  
 PF 17-NOV-2000; 2000WO-US031543.  
 XX  
 PR 19-NOV-1999; 99US-0166336P.  
 PR 29-NOV-1999; 99US-0167785P.  
 PR 08-MAR-2000; 2000US-0187844P.  
 PR 16-NOV-2000; 2000US-00715417.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Lichenstein H, Vernet C, Fernandes E;  
 XX  
 XX WPI; 2001-648134/74.  
 DR N-PSDB; AAF87113.  
 DR  
 XX Novel human polypeptides and the nucleic acids that encode them useful  
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and  
 PT immune disorders.  
 PT  
 XX  
 PS Claim 1; Page 10-13; 141pp; English.  
 XX  
 XX This sequence is the NOV2 protein. The invention relates to the NOV1-  
 CC NOV16 proteins, and their coding sequences. The proteins have Cytostatic;  
 CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular  
 CC activities. The sequences may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate NOVX expression. They  
 CC may be used to treat disorders associated with decreased expression by  
 CC rectifying mutations or deletions in a patient's genome that affect the  
 CC activity of protein by expressing inactive proteins or to supplement the  
 CC patients own production of protein. They are used to produce NOVX  
 CC proteins, by inserting the nucleic acid into a cell and culturing it to  
 CC express the protein. The DNA may be used as DNA probes in assays to  
 CC detect and quantitate the presence of similar DNAs in samples, and which  
 CC patients may need restorative therapy. The NOVX protein may also be used  
 CC as antigens in the production of antibodies (Abs) against NOVX and in  
 CC assays to identify modulators of NOVX expression and activity. The anti-  
 CC NOVX Abs and antagonist are used to down regulate expression and  
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX  
 CC in samples. Disorders that may be prevented, diagnosed and/or treated  
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and  
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of  
 CC serine/threonine kinases, and are used to treat kinase-related disorders  
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).  
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)  
 CC -like super family and are involved in, e.g. regulation of cell  
 CC development, apoptosis, cell adhesion, growth migration, cell structure  
 CC and motility and protein management, and are used to treat cancers,  
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.  
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to  
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT  
 CC syndrome and marfan syndrome  
 XX  
 SQ Sequence 315 AA;

Query Match 16.1%; Score 57; DB 4; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-47;  
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 95 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVGVIGVDPNPKSEMDVFCYRMK 151  
 DB 239 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVGVIGVDPNPKSEMDVFCYRMK 295  
 RESULT 13  
 ID ABG72500  
 XX ABG72500 standard; protein; 93 AA.  
 AC  
 XX  
 XX ABG72500;  
 DT 19-FEB-2003 (first entry)  
 XX  
 XX Human 190kDa Hyaluronan receptor for endocytosis (HARE) LINK domain.

XX Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
 KW tumour; gene therapy; human; LINK domain.  
 XX Homo sapiens.  
 XX WO200286093-A2.  
 XX 31-OCT-2002.  
 XX 25-APR-2002; 2002WO-US013209.  
 XX 25-APR-2001; 2001US-00842930.  
 PR 25-APR-2001; 2001US-0286468P.  
 XX  
 PA (WEIG/) WEIGEL P H.  
 PA (WEIG/) WEIGEL J A.  
 XX  
 PI Weigel PH, Weigel JA;  
 XX  
 XX WPI; 2003-093126/08.  
 DR  
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject  
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
 PT epitope.  
 XX  
 PS Example; Page 77; 167pp; English.  
 XX  
 CC The invention describes a method of targeting a compound to a cell or  
 CC tissue of an individual expressing a functionally active hyaluronan (HA)  
 CC receptor for endocytosis (HARE) or a cell that does not express  
 CC functionally active HARE. The method involves using HA molecule, a  
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
 CC method is useful for targeting a compound, preferably a chemotherapeutic  
 CC agent or a radioisotope to cell of an individual, especially a human,  
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
 CC method useful for preventing interaction between a cell having at least  
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
 CC its surface. This second method is useful for preventing metastasis by  
 CC preventing interaction between tumour cells having HA, CD or CDS coat and  
 CC non-tumour cells expressing HARE on its surface. The invention also  
 CC describes a method useful for detecting the presence of HA, CD and CDS in  
 CC a biological fluid. This sequence represents the LINK domain of the human  
 CC 190kDa Hyaluronan receptor for endocytosis (HARE) that may be present in  
 CC other HARE-like proteins  
 XX  
 SQ Sequence 93 AA;  
 Query Match 15.3%; Score 54; DB 6; Length 93;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-45;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 95 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVGVIGVDPNPKSEMDVFCY 148  
 DB 40 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVGVIGVDPNPKSEMDVFCY 93  
 RESULT 14  
 ID AAM47675  
 XX AAM47675 standard; protein; 1431 AA.  
 XX  
 AC AAM47675;  
 XX  
 XX 22-FEB-2002 (first entry)  
 DT  
 XX Rat Hyaluronic Acid Receptor for Endocytosis, HARE.  
 DE  
 XX HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;  
 KW chondroitin sulphate; extracellular matrix; cartilage; skin;  
 KW vitreous humour; endocytic receptor; glycosaminoglycan; rat.

```

XX OS Rattus norvegicus.
XX PN WC200181544-A2.
XX PD 01-NOV-2001.
XX PF 25-APR-2001; 2001WO-US013403.
XX PR 25-APR-2000; 2000US-0199538P.
XX PR 02-NOV-2000; 2000US-0245320P.
XX XX
XX PA (WEIG/) WEIGEL P A.
XX PA (ZHOU/) ZHOU B.
XX PA (WEIG/) WEIGEL J A.
XX XX
XX PI Weigel PA, Zhou B, Weigel JA;
XX DR WPI; 2002-049271/06.
XX DR N-PSDB; ABA04648.
XX XX
XX PT New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
XX PT identifying agents that inhibit binding to hyaluronic acid, and related
XX PT nucleic acid.
XX PS Claim 20; Fig 21; 263pp; English.
XX XX
XX CC The present invention relates to sequences for rat and human HARE
XX CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, ABA04675
XX CC and ABA047684). HARE can bind specifically to at least one of hyaluronic
XX CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
XX CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
XX CC particular cartilage, skin and vitreous humour. HARE is the endocytic
XX CC receptor responsible for removing HA and other glycosaminoglycans from
XX CC the circulation. The present sequence is rat HARE
XX SQ Sequence 1431 AA;

Query Match 7.1%; Score 25; DB 5; Length 1431;
Best Local Similarity 100.0%; Pred. No. 2.7e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 FQDTTGVFHLRSPGQYKLTFDKA 74
Db 1057 FQDTTGVFHLRSPGQYKLTFDKA 1081
|||||
RESULT 15
ID ABG72498 standard; protein; 1431 AA.
XX AC ABG72498;
XX DT 19-FEB-2003 (first entry)
XX DE
XX DE Rat 175kDa Hyaluronan receptor for endocytosis (HARE).
XX XX
XX KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
XX KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
XX KW tumour; gene therapy; rat; receptor.
XX XX
XX OS Rattus sp.
XX FN WO200286093-A2.
XX PD 31-OCT-2002.
XX XX
XX PF 25-APR-2002; 2002WO-US013209.
XX XX
XX PR 25-APR-2001; 2001US-00842930.
XX PR 25-APR-2001; 2001US-0286468P.
XX XX
XX PA (WEIG/) WEIGEL P H.

```

```

PA (WEIG/) WEIGEL J A.
XX PI Weigel PH, Weigel JA;
XX DR WPI; 2003-093126/08.
XX DR N-PSDB; ABX13821.
XX XX
XX PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
XX PT expressing functional active hyaluronan receptor for endocytosis of HARE,
XX PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
XX PT epitope.
XX XX
XX PS Example; Fig 2; 167pp; English.
XX XX
XX CC The invention describes a method of targeting a compound to a cell or
XX CC tissue of an individual expressing a functionally active hyaluronan (HA)
XX CC receptor for endocytosis (HARE) or a cell that does not express
XX CC functionally active HARE. The method involves using HA molecule, a
XX CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
XX CC a monoclonal antibody raised against a HA-binding domain of HARE. The
XX CC method is useful for targeting a compound, preferably a chemotherapeutic
XX CC agent or a radioisotope to cell of an individual, especially a human,
XX CC expressing HARE on its surface (e.g. gene therapy). Also described is a
XX CC method useful for preventing interaction between a cell having at least
XX CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
XX CC its surface. This second method is useful for preventing metastasis by
XX CC preventing interaction between tumour cells having HA, CD or CDS coat and
XX CC non-tumour cells expressing HARE on its surface. The invention also
XX CC describes a method useful for detecting the presence of HA, CD and CDS in
XX CC a biological fluid. This is the amino acid sequence of the rat 175kDa
XX CC Hyaluronan receptor for endocytosis (HARE)
XX SQ Sequence 1431 AA;

Query Match 7.1%; Score 25; DB 6; Length 1431;
Best Local Similarity 100.0%; Pred. No. 2.7e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 FQDTTGVFHLRSPGQYKLTFDKA 74
Db 1057 FQDTTGVFHLRSPGQYKLTFDKA 1081
|||||
RESULT 16
AAU00893
ID AAU00893 standard; protein; 106 AA.
XX AC AAU00893;
XX DT 04-JUL-2001 (first entry)
XX DE
XX DE Human cancer related protein 28.
XX XX
XX KW Human; cancer related protein; food additive; preservative; immunogen;
XX KW antibody; bone cancer; adrenal cancer; bone marrow cancer; breast cancer;
XX KW gastrointestinal cancer; liver cancer; lung cancer; urogenital cancer;
XX KW immune disorder; Addison's disease; allergy; diabetes mellitus;
XX KW autoimmune haemolytic anaemia; autoimmune thyroiditis; Crohn's disease;
XX KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
XX KW acquired immunodeficiency syndrome; AIDS; cardiovascular disorder;
XX KW myocardial ischaemia; wound healing; neurological disorder;
XX KW Parkinson's disease; Alzheimer's disease; cerebral anoxia; epilepsy;
XX KW viral infection; bacterial infection; fungal infection;
XX KW parasitic infection; agonist; antagonist.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Region 62..70
XX FT Region /label= Immunogenic_epitope
XX FT Region 72..103
XX FT Region /label= Immunogenic_epitope
XX XX

```

PN WO200118014-A1.  
 XX 15-MAR-2001.  
 XX 30-AUG-2000; 2000WO-US023794.  
 XX 03-SEP-1999; 99US-0152296P.  
 XX 06-OCT-1999; 99US-0158003P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Roschke V;  
 XX WPI; 2001-235186/24.  
 XX N-PSDB; AAS00854.  
 XX Twenty nine nucleic acid molecules encoding human cancer associated  
 PT proteins, useful in the prevention, treatment and diagnosis of cancer,  
 PT immune disorders, cardiovascular disorders and neurological diseases.  
 XX Disclosure; Page 403; 427pp; English.  
 XX The sequence represents a novel Human cancer related protein. The  
 CC polynucleotides and polypeptides are useful for preventing, treating or  
 CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,  
 CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities. The polynucleotide are useful for chromosome  
 CC identification. The nucleic acids, protein, antibodies, agonists and  
 CC antagonists are useful in the diagnosis, treatment and prevention of  
 CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital), immune disorders  
 CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired  
 CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as  
 CC myocardial ischaemia, wound healing, neurological diseases (e.g.  
 CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)  
 CC and infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Numerous examples of each type of disorder are given in the  
 CC specification  
 XX Sequence 106 AA;  
 SQ  
 Query Match 5.7%; Score 20; DB 4; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 326 IILVTGAVAAAYSYFRINR 345  
 Db 23 IILVTGAVAAAYSYFRINR 42  
 RESULT 17  
 ABG72512  
 ID ABG72512 standard; protein; 18 AA.  
 XX  
 AC ABG72512;  
 XX  
 DT 19-FEB-2003 (first entry)  
 XX  
 DE Human 190kDa Hyaluronan receptor for endocytosis putative motif #12.  
 XX  
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
 KW tumour; gene therapy; human; motif.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200286093-A2.  
 XX  
 PD 31-OCT-2002.  
 XX

PF 25-APR-2002; 2002WO-US013209.  
 XX  
 PR 25-APR-2001; 2001US-00842930.  
 PR 25-APR-2001; 2001US-0286468P.  
 XX  
 PA (WEIG/) WEIGEL P H.  
 PA (WEIG/) WEIGEL J A.  
 XX  
 PI Weigel PH, Weigel JA;  
 XX WPI; 2003-093126/08.  
 DR  
 XX  
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject  
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
 PT epitope.  
 XX  
 PS Claim 23; Page 77; 167pp; English.  
 XX  
 CC The invention describes a method of targeting a compound to a cell or  
 CC tissue of an individual expressing a functionally active hyaluronan (HA)  
 CC receptor for endocytosis (HARE) or a cell that does not express  
 CC functionally active HARE. The method involves using HA molecule, a  
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
 CC method is useful for targeting a compound, preferably a chemotherapeutic  
 CC agent or a radioisotope to cell of an individual, especially a human,  
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
 CC method useful for preventing interaction between a cell having at least  
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
 CC its surface. This second method is useful for preventing metastasis by  
 CC preventing interaction between tumour cells having HA, CD or CDS coat and  
 CC non-tumour cells expressing HARE on its surface. The invention also  
 CC describes a method useful for detecting the presence of HA, CD and CDS in  
 CC a biological fluid. This sequence represents a putative motif from the  
 CC human 190kDa Hyaluronan receptor for endocytosis (HARE) that may be  
 XX present in other HARE-like proteins  
 SQ Sequence 18 AA;  
 Query Match 5.1%; Score 18; DB 6; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 TGPGRKCKCKSHYVGDG 19  
 Db 1 TGPGRKCKCKSHYVGDG 18  
 RESULT 18  
 AA47702  
 ID AA47702 standard; peptide; 15 AA.  
 XX  
 AC AA47702;  
 XX  
 DT 22-FEB-2002 (first entry)  
 XX  
 DE Human HARE peptide fragment #6.  
 XX  
 KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;  
 KW chondroitin sulphate; extracellular matrix; cartilage; skin; rat;  
 KW vitreous humour; endocytic receptor; glycosaminoglycan.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200181544-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 XX 25-APR-2001; 2001WO-US013403.  
 PF  
 XX 25-APR-2000; 2000US-0199538P.  
 PR  
 PR 02-NOV-2000; 2000US-0245320P.



XX PA (WEIG/) WEIGEL P A.  
 XX PA (ZHOU/) ZHOU B.  
 XX PA (WEIG/) WEIGEL J A.  
 XX PI Weigel PA, Zhou B, Weigel JA;  
 XX WI; 2002-049271/06.  
 XX  
 XX PT New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for  
 PT identifying agents that inhibit binding to hyaluronic acid, and related  
 PT nucleic acid.  
 XX  
 XX PS Example; Page 95; 263pp; English.  
 XX  
 XX CC The present invention relates to sequences for rat and human HARE  
 CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675  
 CC and AAM47684). HARE can bind specifically to at least one of hyaluronic  
 CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin  
 CC sulphate (CS). HA is an extracellular matrix component of all tissues, in  
 CC particular cartilage, skin and vitreous humour. HARE is the endocytic  
 CC receptor responsible for removing HA and other glycosaminoglycans from  
 CC the circulation. The present sequence is a peptide fragment of HARE,  
 CC which was used in an example from the present invention  
 XX  
 XX SQ Sequence 15 AA;  
 Query Match 4.2%; Score 15; DB 5; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 61 RSPGLGQYKLTFFDKAR 75  
 DB 1 RSPGLGQYKLTFFDKAR 15  
 RESULT 19  
 AAM47690  
 ID AAM47690 standard; peptide; 13 AA.  
 AC AAM47690;  
 XX  
 XX DT 22-FEB-2002 (first entry)  
 XX  
 XX DE Human HARE peptide fragment PR 1825 #2.  
 XX  
 XX KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;  
 KW chondroitin sulphate; extracellular matrix; cartilage; skin; human; rat;  
 KW vitreous humour; endocytic receptor; glycosaminoglycan.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN WO200181544-A2.  
 XX  
 XX PD 01-NOV-2001.  
 XX  
 XX PF 25-APR-2001; 2001WO-US013403.  
 XX  
 XX PR 25-APR-2000; 2000US-0199538P.  
 XX  
 XX PR 02-NOV-2000; 2000US-0245320P.  
 XX  
 XX PA (WEIG/) WEIGEL P A.  
 XX PA (ZHOU/) ZHOU B.  
 XX PA (WEIG/) WEIGEL J A.  
 XX PI Weigel PA, Zhou B, Weigel JA;  
 XX WI; 2002-049271/06.  
 XX  
 XX PT New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for  
 PT identifying agents that inhibit binding to hyaluronic acid, and related  
 PT nucleic acid.  
 XX

PS Example; Page 94; 263pp; English.  
 XX  
 XX CC The present invention relates to sequences for rat and human HARE  
 CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675  
 CC and AAM47684). HARE can bind specifically to at least one of hyaluronic  
 CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin  
 CC sulphate (CS). HA is an extracellular matrix component of all tissues, in  
 CC particular cartilage, skin and vitreous humour. HARE is the endocytic  
 CC receptor responsible for removing HA and other glycosaminoglycans from  
 CC the circulation. The present sequence is a peptide fragment of HARE,  
 CC which was used in an example from the present invention  
 XX  
 XX SQ Sequence 13 AA;  
 Query Match 3.7%; Score 13; DB 5; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 213 GTLFVPQNSGLGE 225  
 DB 1 GTLFVPQNSGLGE 13  
 RESULT 20  
 ABG72513  
 ID ABG72513 standard; protein; 13 AA.  
 AC ABG72513;  
 XX  
 XX DT 19-FEB-2003 (first entry)  
 XX  
 XX DE Human 190kDa Hyaluronan receptor for endocytosis putative motif #13.  
 XX  
 XX KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
 KW tumour; gene therapy; human; motif.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN WO200286093-A2.  
 XX  
 XX PD 31-OCT-2002.  
 XX  
 XX PF 25-APR-2002; 2002WO-US013209.  
 XX  
 XX PR 25-APR-2001; 2001US-00842930.  
 XX  
 XX PR 25-APR-2001; 2001US-0286468P.  
 XX  
 XX PA (WEIG/) WEIGEL P H.  
 XX PA (WEIG/) WEIGEL J A.  
 XX PI Weigel PH, Weigel JA;  
 XX WI; 2003-093126/08.  
 XX  
 XX PT Targeting compounds e.g. chemotherapeutic agent to cell of subject  
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
 PT epitope.  
 XX  
 XX PS Claim 23; Page 77; 167pp; English.  
 XX  
 XX CC The invention describes a method of targeting a compound to a cell or  
 CC tissue of an individual expressing a functionally active hyaluronan (HA)  
 CC receptor for endocytosis (HARE) or a cell that does not express  
 CC functionally active HARE. The method involves using HA molecule, a  
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
 CC method is useful for targeting a compound, preferably a chemotherapeutic  
 CC agent or a radioisotope to cell of an individual, especially a human,  
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
 CC method useful for preventing interaction between a cell having at least  
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on

CC its surface. This second method is useful for preventing metastasis by  
 CC preventing interaction between tumour cells having HA, CD or CDS coat and  
 CC non-tumour cells expressing HARE on its surface. The invention also  
 CC describes a method useful for detecting the presence of HA, CD and CDS in  
 CC a biological fluid. This sequence represents a putative motif from the  
 CC human 190Kda Hyaluronan receptor for endocytosis (HARE) that may be  
 CC present in other HARE-like proteins  
 XX

SQ Sequence 13 AA;

Query Match 3.7%; Score 13; DB 6; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 PIDRCLQDNGQCH 40  
 |||||  
 Db 1 PIDRCLQDNGQCH 13

Search completed: April 1, 2004, 09:56:56  
 Job time : 59 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 3238 Seconds  
(without alignments)  
3255.513 Million cell updates/sec

Title: US-09-466-778B-11  
Perfect score: 1865  
Sequence: 1 MTGPGKHCKSHYVDGL.....ALAAYSYFIRNKTIGFXHF 353

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Zgapop 6.0, Zgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2/1/USPTO.spool/US09466778/runat\_01042004\_095642\_10592/app\_query.fasta\_1.519  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09466778 @CGN 1 1 3437 @runat\_01042004\_095642\_10592 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:

- 1: em estba:\*
- 2: em esthum:\*
- 3: em estin:\*
- 4: em estmu:\*
- 5: em estov:\*
- 6: em estpl:\*
- 7: em estro:\*
- 8: em htc:\*
- 9: gb est1:\*
- 10: gb est2:\*
- 11: gb htc:\*
- 12: gb est3:\*
- 13: gb est4:\*
- 14: gb est5:\*
- 15: em estfun:\*
- 16: em estom:\*
- 17: em gss hum:\*
- 18: em gss inv:\*
- 19: em gss pln:\*
- 20: em gss vrt:\*
- 21: em gss fun:\*
- 22: em gss mam:\*
- 23: em gss mus:\*
- 24: em gss pro:\*
- 25: em gss rod:\*
- 26: em gss phg:\*
- 27: em gss vrl:\*
- 28: gb gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1487	79.7	3085	11	AK034522	AK034522 Mus muscu
2	963	51.6	819	13	BY717115	BY717115 BX717115
3	718	38.5	694	14	CA376384	CA376384 654727 NC
4	703	37.7	683	12	BJ523552	BJ523552 BJ523552
5	661.5	35.5	3025	11	BC049247	BC049247 Mus muscu
6	638	34.2	1122	14	CK028027	CK028027 AGENCOURT
7	604.5	32.4	722	13	BY739778	BY739778 BX739778
8	590	31.6	733	12	BI832486	BI832486 603082278
9	575.5	30.9	2368	11	BC019712	BC019712 Mus muscu
10	557.5	29.9	459	14	T47504	T47504 yb14f01.r1
11	552	29.6	3454	11	AK028441	AK028441 Mus muscu
12	504.5	27.1	1201	13	BX446538	BX446538 BX446538
13	462	24.8	572	14	CB268774	CB268774 1007680 H
14	441.5	23.7	865	13	BY705197	BY705197 BX705197
15	441	23.6	719	14	CA342417	CA342417 672284 NC
16	424.5	22.8	761	13	BX861168	BX861168 BX861168
17	418.5	22.4	878	14	CD359001	CD359001 AGENCOURT
18	384.5	20.6	690	10	BF342700	BF342700 602013653
19	384	20.6	962	10	BF300451	BF300451 602032375
20	375	20.1	570	10	BE233224	BE233224 139315 MA
21	356	19.1	557	12	BM258620	BM258620 523466 MA
22	353.5	19.0	798	14	CK030769	CK030769 AGENCOURT
23	350	18.8	721	12	BI219897	BI219897 602936347
24	345.5	18.5	646	12	BM767207	BM767207 K-EST0049
25	344	18.4	895	29	CNS03DSC	AL238989 Tetraodon
26	343.5	18.4	589	12	BM769767	BM769767 K-EST0053
27	339	18.2	522	13	BX866658	BX866658 BX866658
28	335	18.0	388	10	AW435659	AW435659 74622 MAR
29	334	17.9	1253	13	BQ884322	BQ884322 AGENCOURT
30	333.5	17.9	725	14	CA362129	CA362129 636116 NC
31	329.5	17.7	564	12	BM181159	BM181159 fv94h10.Y
32	329.5	17.7	567	12	BM181146	BM181146 fv94g09.Y
33	320.5	17.2	928	12	BI906024	BI906024 603062355
34	316	16.9	804	12	BI144649	BI144649 602909939
35	313.5	16.8	467	12	BI132408	BI132408 AR030G05L
36	309.5	16.6	971	29	CNS03XY8	AL265625 Tetraodon
37	304.5	16.3	429	10	BF870695	BF870695 IL3-ET011
38	296.5	15.9	439	12	BM694227	BM694227 UI-E-C11-
39	293.5	15.7	482	12	BM769766	BM769766 K-EST0053
40	289.5	15.5	704	14	CA364659	CA364659 639611 NC
41	285.5	15.3	520	12	BG712246	BG712246 pglin.pk0
42	282	15.1	603	13	BQ131025	BQ131025 fz47905.Y
43	282	15.1	604	10	AW595426	AW595426 fk37902.Y
44	281	15.1	789	12	BG935766	BG935766 SSI-0071
45	280	15.0	606	14	CA347334	CA347334 678437 NC

## ALIGNMENTS

RESULT 1  
AK034522  
LOCUS  
DEFINITION  
AK034522 3085 bp mRNA linear HTC 18-SEP-2003  
Mus musculus adult male diencephalon cDNA, RIKEN full-length  
enriched library, clone:9330210123 product:similar to CD44-LIKE  
PRECURSOR FELL [Homo sapiens], full insert sequence.  
ACCESSION  
AK034522  
VERSION  
AK034522.1 GI:26330004  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
PUBMED 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
PUBMED 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Todaga, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
PUBMED 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5

6 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3085)

Adachi, J., Aizawa, K., Akimura, T., Atakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saito, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

FEATURES  
source  
1..3085  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:9330210123"

Db 1816 AAGAAATGT--GCAACATTGTGGGATTGTAGACTACGGAACCGAGACCAACAGAGT 1872

Qy 141 GluMetTrrpAapValPheCysTyrArgMetLysAspValAenCysThr\*\*\*LysValGly 160

Db 1873 GAATGTGGGACGCTCTCTGTACAGATGAAGATGTTAACTGCACCTGCAAGCGCAGG 1932

Qy 161 TyrValGlyAapGlyPheSerTyrSerGlyAenLeuGlnValLeuMetSerPhePro 180

Db 1933 TACGTGGGCGATGCTCTCTGCAATGGGAACCTGTGCAAGTCCTCATGCTCCTCC 1992

Qy 181 SerLeuThrAenPheLeuThrGluValLeuAlaTyrSerAenSerSerAlaArgGlyArg 200

Db 1993 TCATCTACGAATCTCTGACAGAGTGCTGCTTTTCCAGGAGCTCAGCCCAAGGCGCG 2052

Qy 201 AlaPheLeuGluHisLeuThrAspLysSerIleArgGlyThrLeuPheValProGlnAsn 220

Db 2053 GCGTTTGTGAACACCTGACTGCTGCTCAATGATGCGACCCCTGTTGTGCCACAGAC 2112

Qy 221 SerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240

Db 2113 AGTGGGCTACCGAAAAATAGAGGCTGTCTGGGCGGACATTGAGCACCACTCACTAAT 2172

Qy 241 ValSerMetPhePheTyrAsnAspLeuValAenGlyThrThrLeuGlnThrArgLeuGly 260

Db 2173 GTCACAGTCTCCTTTACGATGACCTGTGCAATGATGCTGCTGAGAGCTAGGCTGGGA 2232

Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280

Db 2233 AGCCAACCTGCTCATCCTCCAGCAGCAGCAGCAGCTCCAC---CAAGAGCGCAGGTTCGTG 2289

Qy 281 AspGlyArgAspThrLeuGluThrPaspIleCysAlaSerAenGlyIleThrHisValIle 300

Db 2290 GATGGAAGAGCAATCTCTGACGTGGAGACATCATGCTCTTAACGGGGTTCCTCATCAT 2349

Qy 301 SerArg\*\*\*LeuLysAlaProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320

Db 2350 TCTGAACCTTTGAAGCTCTCTCCAGCGCGCAGCGCTGCCACTCTGGCTTAGGAACA 2409

Qy 321 GlyIlePhe\*\*\*\*\*IleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340

Db 2410 GGCATATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2469

Qy 341 PheArgIleAenArgLysThrIleGlyPhe\*\*\*HisPhe 353

Db 2470 TTCGGCTAAACAGAGAACCACTGGCTTCGGGGGCTTT 2508

## RESULT 2

EX717115 819 bp mRNA linear EST 18-NOV-2003

LOCUS EX717115 XGC-tadpole Silurana tropicalis cDNA clone TtpA026k09 5',

DEFINITION mRNA sequence.

ACCESSION EX717115 GI:38389795

VERSION EX717115

KEYWORDS EST.

SOURCE Silurana tropicalis (western clawed frog)

ORGANISM Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Silurana.

1 (bases 1 to 819)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11\_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TtpA026k09.p1kSP6

Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Nigel Garrett.

cDNA was oligo dt primed from sug of poly A+ RNA from tadpole

embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli DH10B.

FEATURES

Location/Qualifiers

source

1..819

/organism="Silurana tropicalis"

/mol\_type="mRNA"

/db\_xref="taxon:8364"

/clone="TtpA026k09"

/dev\_stage="tadpole (stage 35-40)"

/lab\_host="E. coli DH10B"

/clone\_lib="XGC-tadpole"

/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from sug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

## ORIGIN

Alignment Scores:

Pred. No.: 5,81e-107 Length: 819

Score: 963.00 Matches: 179

Percent Similarity: 79.12% Conservative: 37

Best Local Similarity: 65.57% Mismatches: 51

Query Match: 13 Indels: 6

DB: 1 Gaps: 1

US-09-466-778b-11 (1-353) x BX717115 (1-819)

Qy 15 TyrValGlyAapGlyLeuAenCysGluProGluGlnLeuProIleAspArgCysLeuGln 34

Db 1 TATATTGGCATGGGTTGAGTTGTAAGTGAAGAAAGCTTCCAATTAATCGCTGCCTCAA 60

Qy 35 AspAsnGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGlnAspThr 54

Db 61 GACAAAGGCAATGCCATGCTGATGATTAATGATCTACACTTACAGATTCCACC 120

Qy 55 ValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAla 74

Db 121 GTCGGTGTTCATTTGCGTTCTCCAAAGGGGTGTACAAATACAACTACAGCGAAGCC 180

Qy 75 ArgGluAlaCysAlaAenGluAlaAlaThrMetAlaThrTyrAenGlnLeuSerTyr\*\*\* 94

Db 181 ATGAAGCTTTGTAATGATGAGACTGCAACCAATAGCAACCTCAATAGCTGTCTATGCA 240

Qy 95 GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrrpLeuGluThrGlyArgValAlaTyr 114

Db 241 CAACAGCGGGATTCCATCTCTGTGAGCTGGTGGTGGATGGCTTAGAGTCGCATAT 300

Qy 115 ProThrAlaPheAlaSerGlnAenCysGlySerGlyValValGlyIleValAspTyrGly 134

Db 301 CCTACCACTTATTCCAAACCAAGCTGTGGATCAGGATTTGTTGAATAGTACACTATGA 360

Qy 135 ProArgProAenLysSerGluMetTrrpAspValPheCysTyrArgMetLysAspValAsn 154

Db 361 CCCCGTGTCAACCTTAAGTACAGCTGAGATGTTTCTGCTACAGATTAAGATGTGTCA 420

Qy 155 CysThr\*\*\*LysValGlyTyrValGlyAspGlyPheSerTyrSerGlyAenLeuGln 174

Db 421 TGTATTGCAACCTGGCTATGTGGAGATGGCTACACTTGTAAATGGAACCTCTGCAA 480

Qy 175 ValLeuMetSerPheProSerLeuThrAenPheLeuThrGluValLeuAlaTyrSerAsn 194

Db 481 GTCCTGACATCTTTTCCTTCCTTTCAAAATTCCTTACGGAAATTTTGGTGTACTCTAAC 540

Qy 195 SerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThr 214

Db 541 ACTTCAGTGAAGGGAAGAATTTTCACTACCTCACTAATGAAAATCAGACCTGTCTGGACGGATATA 600

Qy 215 LeuPheValProGlnAenSerGlyLysGlyValGluAenGlnThrLeuSerGlyArgAspIle 234

Db 601 CTGTTTGGCCCAAGCAATGATGGACTTAATGAAAATCAGACCTGTCTGGACGGATATA 660

```

QY 235 GluHisLeuAlaAsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThr 254
      |||...|||
DB 661 GAATATCATCTGGCAATGTCAGCATGTTCTTTGAGGACCTGAGCAATGGAACAACT 720
      |||...|||
QY 255 LeuGlnThrArgLeuGlySerLysLeuLeuThrAspArgGlnAspProLeuHisPro 274
      |||...|||
DB 721 CTTCAACAGCAATAGGCGCACAAAGTATTAAATTTCTTTTGAATGACCCAGCCAGCAAA 780
      |||...|||
QY 275 -----ThrGluThrArgCysValAsp 281
      |||...|||
DB 781 ATAACCACTGATTCATGACTCCAAACCATGATGTTGGAT 819

RESULT 3
LOCUS CA376384 694 bp mRNA linear EST 06-NOV-2002
DEFINITION 654727 NCCOWA 1RT Oncorhynchus mykiss cDNA clone lRT35N18_D_G09 5',
      mRNA sequence.
ACCESSION CA376384
VERSION CA376384.1 GI:24694516
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
      Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 694)
      Rexroad,C.E. and Keele,J.W.
      Sequence analysis of a rainbow trout normalized cDNA library
      Unpublished (2002)
      Contact: Rexroad CE
      USDA, ARS, National Center for Cool and Cold Water Aquaculture
      11876 Leetown Road, Kearneysville, WV 25430, USA
      Tel: 304 724 8340 x2129
      Fax: 304 725 0351
      Email: crexroad@cccwa.ars.usda.gov
      Single pass sequencing. Bases called with phred v0.020425.c and
      trimmed with the aid of the trim_alt option. Vector identified by
      cross_match v0.990329.
      Seq primer: AGCGGATACAAATTTTCACACAGGA.
FEATURES
      source
      1..694
      /organism="Oncorhynchus mykiss"
      /mol_type="mRNA"
      /db_xref="taxon:8022"
      /clone="lRT35N18_D_G09"
      /tissue_type="pooled"
      /lab_host="DH10B"
      /clone_lib="NCCOWA 1RT"
      /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
      Library made from pooled tissue from brain, gill, liver,
      spleen, muscle, and kidney."
ORIGIN
Alignment Scores: 5.32e-77 Length: 694
Pred. No.: 718.00 Matches: 126
Score: 718.00 Mismatches: 45
Percent Similarity: 74.67% Conservative: 126
Best Local Similarity: 55.02% Mismatches: 58
Query Match: 38.50% Indels: 0
DB: 14 Gaps: 0

US-09-466-778b-11 (1-353) x CA376384 (1-694)

QY 25 GluGlnLeuProfileAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLys 44
      |||...|||
DB 4 AAGAGCTTCTTGACACCTGTTCTTTCAGGAGAACGGGACGTGCCACTGACGCCAG 63
      |||...|||
QY 45 CysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeu 64
      |||...|||
DB 64 TGTACCGACCTTCACTACGAGGACAGCACAGTCGGTGTGTTCCACTTCGGTCCCAAG 123
      |||...|||
QY 65 GlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaThr 84
      |||...|||

```

```

DB 124 GGACAGTACAGCTGAACCTATACATCGGCCCCAGGAGGCTCCACAGAGAGGGGGGAAACC 183
      |||...|||
QY 85 MetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAla 104
      |||...|||
DB 184 ATCCGCCAGTACACACAGCTCTCTCCGACAGCAGGCTGGGTTCAACCTGTGTGGCGCT 243
      |||...|||
QY 105 GlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGly 124
      |||...|||
DB 244 AGCTGGCTGGACCCAGGCGCAGAGTGGCTTACCCACCACCTACTCTCAACCCCAAGTGTGGC 303
      |||...|||
QY 125 SerGlyValValGlyValLeuValAspTyrGlyProArgProAsnLysSerGluMetTrpAsp 144
      |||...|||
DB 304 TTTGGTCACGTGGGATCGTAGACTACGGCTTACGCAACACCTCAGTAGAGACCTGGGAC 363
      |||...|||
QY 145 ValPheCysTyrArgMetLysAspValAsnCysThr***LysValGlyTyrValGlyAsp 164
      |||...|||
DB 364 ACCTTCTGCTACAGGCTCAAGAGGTGAATGTGAATGAAGCAGGTTATATAGGAGAT 423
      |||...|||
QY 165 GlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSerLeuThrAsn 184
      |||...|||
DB 424 GGATACTCCTGCACTGGAAACCTCTCGCAGGTTCTCACTGAGAAGCCACCTCTCCAAC 483
      |||...|||
QY 185 PheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGlu 204
      |||...|||
DB 484 TTCCTCTCTCAATCTGAACTCTCCAGATGTCCTGTCAGGAAACAGTTTATGATG 543
      |||...|||
QY 205 HisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGly 224
      |||...|||
DB 544 CGCCTCAGTAACCTGACCATCCAGTCCACTCTGTTGTACCTGATATATACCGGACTGTAC 603
      |||...|||
QY 225 GluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhe 244
      |||...|||
DB 604 CAGAACCCAGCAGTACCCACAGGACATGAGTACCATCTCTCAGAGGGCGGGCCCTTG 663
      |||...|||
QY 245 PheTyrAsnAspLeuValAsnGlyThr 253
      |||...|||
DB 664 GCCCTGAAAGACCTGACCAACGGCAGC 690
      |||...|||

RESULT 4
LOCUS BJ523552 683 bp mRNA linear EST 09-AUG-2002
DEFINITION BJ523552 MF01SSB cDNA Oryzias latipes cDNA clone MF01SSB002F09 5',
      mRNA sequence.
ACCESSION BJ523552
VERSION BJ523552.1 GI:22182364
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
      Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
      Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE 1 (bases 1 to 683)
      Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
      Medaka EST Project in Takeda's lab
      Unpublished (2001)
      Contact: Tadasu Shin-i
      Center For Genetic Resource Information
      National Institute of Genetics
      1111 Yata, Mishima, Shizuoka 411-8540, Japan
      Tel: 81-559-81-6856
      Fax: 81-559-81-6855
      Email: tshini@genes.nig.ac.jp.
      Location/Qualifiers
      1..683
      /organism="Oryzias latipes"
      /mol_type="mRNA"
      /strain="Hd-rR"
      /db_xref="taxon:8090"
      /clone="MF01SSB002F09"
      /sex="mixture of female and male"
      /tissue_type="whole embryo"
      /dev_stage="segmentation stage 20 - 25"
FEATURES
      source

```

/clone\_lib="MF01SSB cdna"

## ORIGIN

## Alignment Scores:

Pred. No.: 3,59e-75 Length: 683  
 Score: 703.00 Matches: 123  
 Percent Similarity: 73.73% Conservative: 37  
 Best Local Similarity: 56.68% Mismatches: 57  
 Query Match: 37.69% Indels: 0  
 DB: 12 Gaps: 0

US-09-466-778B-11 (1-353) x BJ523552 (1-683)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 Db 32 ATGACGGCTCGGGAAAAAGAACTGGCCCTGTGAAGCACTACCTTGGAGATGGCGTG 91  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAnglyGlnCysHis 40  
 Db 92 AGTGGCACTCATGACAGATCCCGTGTAGCGCTGTCTCAGGCAATGGGGGTGCCAC 151  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 152 TCAGATGCCAGATGTTCTGATCTCCACTTGAAGATAAAGTGTGTGTTCAATGTC 211  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 212 CGCTCCAAACCGGTGAGTAACTACCTATGCTGACGCTCAGCATCGATGTCATCATA 271  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*GlnLysAlaLysTyrHis 100  
 Db 272 GAGGAGGGTCCTTGGCCAGCGCTCCAGCTTCCAGTCTAGCAGGGGGGCTTGAAC 331  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 332 ATGTGTGCTGCTGCTGCTGGACCAAGCGAGTTCATACCCCACTACCTACTCTPAAC 391  
 QY 121 GlnAsnCysGlySerGlyValValGlyValValAspTyrGlyProArgProAsnLysSer 140  
 Db 392 CCCAATGTGGCTTGGACATGGGGTATGTGGACTACGGGTTCGCCAAACACCTGAGC 451  
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*LysValGly 160  
 Db 452 GAGACCTGGGACGGCTTCTGTACCGGATGAAGATGTGAATGTCAATTAACCTGGT 511  
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180  
 Db 512 TATATCGAGACGGTGTGTGGCTGTGTGTGTAACCTGATGAGGTCTCTTAAGTCCAGCCC 571  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 Db 572 AGCTTTCTAACTTCCTCACAAATCCTAAATTAATCTCCAGACGCTCCNAGTCCGGCCGC 631  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheVal 217  
 Db 632 CAGTTGTCAACGCGCTCGGTGACCTCGAGTTCAGTTCAGTCCACCTGTGTGA 582

## RESULT 5

BC049247 3025 bp mRNA linear HTC 28-MAR-2003  
 LOCUS Mus musculus, Similar to stabilin 1, clone IMAGE:5038209, mRNA.  
 DEFINITION BC049247  
 ACCESSION BC049247  
 VERSION BC049247.1 GI:29351595  
 KEYWORDS HTC.

## SOURCE

Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3025)

## REFERENCE

STRAUSBERG,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-MAR-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 99 Row: 9 Column: 14  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis

This clone has the following problem: frame shifted.

## FEATURES

## source

Location/Qualifiers

1..3025

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:5038209"

/tissue\_type="Mammary tumor metastatized to lung. Tumor

arose spontaneously from a senescent normal mammary

(clonal) outgrowth infected with the virus MMTV."

/clone\_lib="NCI CGAP\_Lu29"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

## ORIGIN

## Alignment Scores:

Pred. No.: 5,75e-69 Length: 3025  
 Score: 661.50 Matches: 146  
 Percent Similarity: 52.73% Conservative: 47  
 Best Local Similarity: 39.89% Mismatches: 146  
 Query Match: 35.47% Indels: 27  
 DB: 11 Gaps: 8

US-09-466-778B-11 (1-353) x BC049247 (1-3025)

QY 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21

Db 1612 ACCGGTCCGAACACACACGGCGCTGTGAATGCCAGTAGGCTACGTGGGTGTCAG 1671

QY 22 Cys--GluProGluGlnLeuProIleAspArgCysLeuGlnAspAnglyGlnCysHis 40

Db 1672 TGTCTAGAGAGAGCTTGACCCCTGTGGACAGATGCTTGGAGAGATCATCATTGCCAC 1731

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60

Db 1732 ACTGATGCTTTGTGCACCTGACCTACATTTCCAGGAAACAGCGTGTCTTCCACATC 1791

QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80

Db 1792 CAGGCCACCCAGTGGCGCTTATGCTCTGACCTTCTCAGAGGCCACGAGCGTGTAGGGC 1851

QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*GlnLysAlaLysTyrHis 100

Db 1852 CAGGAGCGCTCTGCTTCACTCCCTCAACTCTCTGCGGCCAACAGCAGTAGTTTCAT 1911

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: gcapbs-i@mail.nih.gov  
Tissue Procurement: Len Zon, Harvard  
cDNA Library Preparation: Open Biosystems  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LIA414828 row: b column: 03  
High quality sequence stop: 751.  
Location/Qualifiers  
1. 1122  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:7053029"  
/tissue\_type="whole body"  
/lab\_host="DH10B"  
/clone\_lib="NIH ZGC 7"  
/note="Vector: pExpress; Site 1: NotI; Site 2: EcoRV;  
Bulk tissue was collected from a whole adult individual  
from the Tuebingen strain. 1st strand cDNA was primed with  
a Not I - oligo(dT) primer, double-stranded cDNA was  
cloned into the Not I and EcoRV sites of pExpress-1.  
Library was size-selected for >1 kb fragments and  
normalized. A non-normalized version of this library is  
also available (NIH ZGC 10). Library was constructed by  
Open Biosystems (Huntsville, AL)"

FEATURES  
source

101 LeuCySerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
1912 GTCTGCTTCGCGGCTGGTGGCAACGGCTCTGCTGCCACCTGTGTTCACCCAGCG 1971  
121 GlnAenCysGlySerGlyValValGlyLeuValAspTyrGlyProArgProAsnLysSer 140  
1972 GCAGACTGTGTAATAATCGTGTAGGTAGTATGACCTGGGGCCGTAAGAACTCTTCG 2031  
141 GluMetTrpAspValPheCysTyrArgMetLysAspValAenCysThr\*\*LysValGly 160  
2032 GAGCTCTGGATGCTACTGATTCGCGTGCAGACGCTGGCTTGCAGTGTGCGGCGCGC 2091  
161 TyrValGlyAspGlyPheSer---TyrSerGlyAenLeuLeuValLeuMetSerPhe 179  
2092 TTCGTGGGTGACGGGATCAGACGCTGCAACGGGAACCTGCTTGTATGCTTGGCGCCACT 2151  
180 ProSerLeuThrAenPheLeuThrGluValLeuAlaTyrSerAenSerSerAlaArgGly 199  
2152 GCCAACTTCTCCACCTTCTATGGGATGCTGTGGGCTATGCCAATGCCACCGAGAGGT 2211  
200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219  
2212 CTGATTTATGATTTCTTCTGAGGATGAGCTACCTACAGACACTCTTCTGTTCTCTGTC 2271  
220 AenSerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisLeuAla 239  
2272 AACAAAGGCTTGTGGACAAATGACGCTGAGTGGCCAGATCTAGAACTCCACGCTTCT 2331  
240 AenValSerMetPhePheTyrAenAspLeuValAenGlyThrLeuGlnThrArgLeu 259  
2332 AATGCCACTTCTTCTGAGTGTCAAT---GCCACCGGGGGACATTCCTCTCCCTGCCACTCA 2388  
260 GlySerLysLeuLeuThrAspArg-----GlnAspProLeuHisProThrGlu 276  
2389 GGTCTTAGCTCTTCATAGTGCATGAGTGGCCCTGACACACTCTCTGTTCTCTGGCC 2448  
277 ThrArgCysValAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAenGlyIle 296  
2449 CCGGGGGCAGTGTGGTTAGCCACGCTCATCGTGGGACATCATGCTTTTCAACGGCATC 2508  
297 ThrHisValIleSerArg\*\*LeuLysAlaPro----- 307  
2509 ATTATGCTCTGCGCAGCCCTCTCATGCTCCCTCCACACTAGGCGAGTGTGGATCT 2568  
308 ---ProAlaProValThrLeu\*\*HisThrGlyLeuGly\*\*GlyIlePhe\*\*\*\*\*Ile 326  
2569 GAGCTCCACCTGTGGCATT-----AGCTGGGGGTGTGTAACCTTCTGGAACA 2619  
327 Ile-----LeuValThrGlyAlaValAlaLeuAlaTyrSerTyrPheArgIleAen 344  
2620 CTGCTGGGCTGTGGCGGAGCTCTC-----TACCTGGTGGCCGA 2661  
345 ArgLysThrIleGlyPhe 350  
2662 GGCAAACTCCAGGTTTC 2679  
RESULT 6  
CK028027 1122 bp mRNA linear EST 26-NOV-2003  
LOCUS AGENCOURT\_16620198 NIH\_ZGC\_7 Danio rerio cDNA clone IMAGE:7053029  
DEFINITION 5', mRNA sequence.  
ACCESSION CK028027  
VERSION CK028027.1 GI:38553951  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 1122)  
REFERENCE NTH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

ORIGIN

Alignment Scores:  
Pred. No.: 8.07e-67 Length: 1122  
Score: 638.00 Matches: 132  
Percent Similarity: 59.75% Conservative: 56  
Best local Similarity: 41.25% Mismatches: 117  
Query Match: 34.21% Indels: 16  
DB: 14 Gaps: 5  
US-09-466-778B-11 (1-353) x CK028027 (1-1122)  
QY 3 GlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAenCys 22  
DB 2 GGACCCAAACGAGCGCGCTGTGAGTGTCTCGATGGCTATGTTGGGAATGGAGTCCAGTGT 61  
QY 23 ---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAenGlyGlnCysHisAla 41  
DB 62 TTGGAGAAAGTGGTCTCTCTCTCTGATCGATCGCTGGAGGATAACGGCGCTGCCACCA 121  
QY 42 AspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArg 61  
DB 122 AAGGGGATATGACAGAGACCTTCATTTCCACACAAAACATGCGAGGTGTTTCCATCTGCGC 181  
QY 62 SerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAenGlu 81  
DB 182 TCACCTGAGGGAAGATATAAATGAATTCACCGAAGCAAGCAGCTTGTGAACAGAG 241  
QY 82 AlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHisLeu 101  
DB 242 GGAGCCATCTTGGCCACGTTTTCCAGCTCTCTGATGCCCAACAGTAGGAATGCACCTG 301  
QY 102 CysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGln 121  
DB 302 TGTGTGGCGGATGATGAGCGGGAAGAGAGGCTATCCGATCCGCTTCTCTCAGCT 361  
QY 122 AenCysGlySerGlyValValGlyLeuValAspTyrGlyProArgProAsnLysSerGlu 141  
DB 362 AAATGGCGGTGATTAATCATGTCGGTATTGGTGTACAAAGATCCAGTCACTCCAGTTGG 421



```

QY 142 MetTrpAspValPheCysTyrArgMetLysAspValAsnValAsnCysThr***LysValGlyTyr 161
Db 422 CTGTACATGCTGCTACTGCTACAGATGAATGAGTGGCTTGTGAATGTGGAGCTGGATAT 481
QY 162 ValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSer 181
Db 482 GTTGGAGCGGGCATTTTGTATGGAGACTTGGCTTGGTGTGTGGCCAACTTCCAAAT 541
QY 182 LeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla 201
Db 542 TTTTCTGTGTTTACACGAATTTACTGAAGTACGACAGGATGGAGAGGACGCAAGACC 601
QY 202 PheLeuGlnHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
Db 602 CTGCTGAAGGCCCTGTCTGCAATCAATTAACATGACCACTTTTATACCCCAAGAACAC 661
QY 222 GlyLeuGlyGluAsnGlnThrLeuSerGlyArgAspIleGluHisLeuAlaAsnVal 241
Db 662 GGATTTGGAGNCACAGACTTTATCTGCTGGAGATTTGCAATATCATATCTCAGCGTC 721
QY 242 ---SerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 722 AACAAATATTCACTATTATGAGAACCTGAAGGCACACACTGCAATCCCGTCTCGCTGGGC 781
QY 261 SerIleLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 782 TC-AAACTCATATAACCACTCCCTCCATTTCCA-----CCGGTCGACACTGAGGAGCGTC 834
QY 281 AspGlyArgAspThrLeuGluTyrAspIle-----CysAla 292
Db 835 AACCTCAGA-----ACCTGGGNATTAAGAGGTATTTTGGCTGAACCATTCGGCC 885
QY 293 SerAsnGlyThrHisValIleSerArg***LeuLysAlaProProAlaProValThr 312
Db 886 ATCAATGGACTGATCCAGCTTCATGATGGGCTTTCAAACTCCCTCGTGGCTTTC 945

```

```

RESULT 7
BX739778 722 bp mRNA linear EST 18-NOV-2003
LOCUS BX739778 XGC-tadpole silurana tropicalis cDNA clone TTPA048d24 5',
DEFINITION mRNA sequence.
ACCESSION BX739778
VERSION 1
KEYWORDS
SOURCE

```

```

ORGANISM
Silurana tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 722)
REFERENCE
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TTPA048d24.plkasP6
Sequencing primer: Sp6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from Sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..722
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA048d24"

```

```

FEATURES
source

```

```

/dev stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4,92e-63 Length: 722
Score: 604.50 Matches: 119
Percent Similarity: 71.22% Conservative: 27
Best Local Similarity: 58.05% Mismatches: 52
Query Match: 32.41% Indels: 7
DB: 13 Gaps: 2
US-09-466-778b-11 (1-353) x BX739778 (1-722)

```

```

QY 152 AspValAsnCysThr***LysValGlyTyrValGlyAspGlyPheSerTyrSerGlyAsn 171
Db 1 GATGTGTGATGATTTGCAAACTCCCTTTGTGGAGATGGCTACACTTGAATGAAGAAC 60
QY 172 LeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGluValLeuAla 191
Db 61 CTGCTCAAGTCTTGACATCTTTTCCTCTCTTCAAAATTCCTTACGGAAATTTTGGTG 120
QY 192 TyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAspLeuSerIle 211
Db 121 TACTCTAACACTTCACTGAAGAGGAAAGAAATTTTCAACTCACTCACTCACTTATCAGTG 180
QY 212 ArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGluAsnGlnThrLeuSerGly 231
Db 181 CAAGCGACTCTGTTGCCCCCAAGCAATGATGGACTTAATGAAATCAGACCCCTGCTGGA 240
QY 232 ArgAspIleGluHisLeuAlaAsnValSerMetPhePheTyrAsnAspLeuValAsn 251
Db 241 CGCGATATAGAAATCATCTGCAAAATGTGAGCATGTTCTTCTTGGAGCACTGAGCAAT 300
QY 252 GlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuThrAspArgGlnAspPro 271
Db 301 GGAACAACTCTTCAAAACCAAGATAGGCCACAGTATTATTAATTTCTTTGACAAATGACCCA 360
QY 272 LeuHisPro-----ThrGluThrArgCysValAspGlyArgAspThr 285
Db 361 GCCAGCAAAATAACCACTGATTCCTCATGCTCCCAACAGATATGTGGATGGAAACCTATT 420
QY 286 LeuGluTyrAspIleCysAsnSerAsnGlyIleThrHisValIleSerArg***LeuLys 305
Db 421 TTGCAGTGGGACATTATTCGTTCCAAATGGAATCATTCACAAATTTGCTGAACCTCTAACT 480
QY 306 AlaProProAlaProValThrLeu***HisThrGlyLeuGly***GlyIlePhe***** 325
Db 481 GCGCTCCAGACCACTAGCTCTG---CATGCTGGCCATGGAGCAGGAATATTTCTTTGGC 537
QY 326 IleIleLeuValThrGlyValAlaLeuAlaTyrSerTyrPheArgIleAsnArg 345
Db 538 ATTGTCTGATCGTGGGCTCTCGCACTCGCAGTATACTATTATAAGAAAGTTCAACAGA 597
QY 346 LysThrIleGlyPhe 350
Db 598 AAAAGATTCCCAATTC 612

```

```

RESULT 8
BX732486 793 bp mRNA linear EST 04-OCT-2001
LOCUS BX732486 Homo sapiens cDNA clone IMAGE:5221410 5',
DEFINITION mRNA sequence.
ACCESSION BX732486
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 793) NIH-MGC http://mgi.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: csapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Inyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11556 row: 1 column: 19 High quality sequence stop: 783. Location/Qualifiers 1..793 /organism="Homo sapiens" /mol\_type="mRNA" /db\_xref="taxon:9606" /clone="IMAGE:5221410" /lab\_host="DH10B" /clone\_lib="NIH\_MGC\_120" /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:  
Pred. No.: 3,496-61 Length: 793  
Score: 590.00 Matches: 127  
Percent Similarity: 63.68% Conservative: 1  
Best Local Similarity: 63.18% Mismatches: 8  
Query Match: 31.64% Indels: 66  
DB: 12 Gaps: 1

US-09-466-778b-11 (1-353) x BI832486 (1-793)

QY 91 LeuSerTyr\*\*\*GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGly 110  
DB 193 CTGCTGTATCCCTCAGGCGCAAGTACCACCTGTCTCAGCAGGCTGGCTGGAGACCGGG 252

QY 111 ArgValAlaTyrProThrAlaPheAlaSerGlnAsnGlySerGlyValValGlyIle 130  
DB 253 CGGGTGGCTACCCACAGCCTTCGCCCTCCAGAACTGTGGCTCTGTGGTGGGATA 312

QY 131 ValAspTyrGlyProArgProAsnLysSerGluMetTrpAspValPheCysTyrArgMet 150  
DB 313 GTGGACTATGGACCTAGACCCCAAGAGTGAATGTGGAGTCTCTCTGCTATCGGATG 372

QY 151 Lys-AspValAsnGlyThr\*\*\*LysValGlyTyrValGlyAspGlyPheSerTyrSerG1 170  
DB 373 AACAGATGTGAATGCACCTGCAAGGTGGGCTATGTGGAGATGGCTTCTCATGCAAGTGG 432

QY 170 YAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThr----- 187  
DB 433 GAACCTGTGTCAGGTCTCTGATGCTTCCCTCACTCAAACTTCTGACGGATTCACT 492

QY 187 ----- 187

DB 493 GAGGAAGAAAGTCTCTCAAACTCTCCACGGCGGAGGAGAGACAGACAGCAAGGAGC 552

QY 187 ----- 187

DB 553 AGCAGCTGAGTTGGCTCACCATGATATTCACAGATGGGTGATGCTTCCCTGGACCTTC 612

QY 187 ----- 187

DB 613 TTCAGCCACCCAGGTCCCTGCACACAGGCGTGGCCCTGTGATGACCTCTTCTTCTCC 672

QY 188 ----GluValLeuAlaTyrSerAsnSerAlaArgGlyArgAlaPheLeuGluHisLe 206  
DB 673 CCAGGAAGTGTGGCTTATTCACACAGCTCAGCTCAGAGGCGTGCATTTCTAGACACT 732

QY 206 uThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGluAs 226  
DB 733 GACTGACCTGTCCATCGCGGACCTC-TTTGTGCCACAGACAGTGGCTGGGGAGAA 791

QY 226 n 226  
DB 792 T 792

RESULT 9  
BC019712 2368 bp mRNA linear HTC 20-SEP-2002  
LOCUS Mus musculus, Similar to stabilin 1, clone IMAGE:3663004, mRNA.  
DEFINITION BC019712  
ACCESSION BC019712  
VERSION BC019712.1 GI:22507492  
KEYWORDS HTC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 2368)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: csapbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: amg@bcm.tmc.edu  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: TRAK Plate: 40 Row: d Column: 23  
This clone has the following problem: frame shifted.

FEATURES  
source 1..2368  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3663004"  
/tissue\_type="Mammary tumor metastasized to lung."  
MRTV-LTR/Wnt1 model. Expression driven by an MRTV-LTR enhancer."  
/clone\_lib="NCI CGAP\_Lu30"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:  
Pred. No.: 1,376-58 Length: 2368  
Score: 575.50 Matches: 147  
Percent Similarity: 43.72% Conservative: 48

Best Local Similarity: 32.96% Mismatches: 144  
 Query Match: 30.86% Indels: 109  
 DB: 11 Gaps: 11

US-09-466-778B-11 (1-353) x BC019712 (1-2368)

QY 2 ThrGlyProGlyHisHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21  
 DB 725 ACCGGTCCGAAACACACGCGCGTGTGAATGCCACGTAGCGTACGTGGGTGATGGGCTCAG 784  
 QY 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 DB 785 TGCTAGAGGAGCTTGAACCCCTGTGGACAGATGCTTGGAGGATCATCACCTTGGCCAC 844  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 DB 845 ACTGATGCTTTGTGCACTGACCTACATTTCCAGGAAAAACAGGCTGGTGTCTCCACATC 904  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 DB 905 CAGGCCACCAAGTGGCCCTTATGGTCTGACCTTCACAGAGCCACGGAAGCGTGTGAGGGC 964  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*GlnLysAlaLysTyrHis 100  
 DB 965 CAGGAGCCGCTTGTCTCTCACTCTCTGCGGCCCAACAGCTAGGTTTTCAT 1024  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 DB 1025 GTCTGCTTGGGGCTGGTGGCAACGCTCTGCTGCCACCTGTGTACCCCGACG 1084  
 QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140  
 DB 1085 GCAGCTGTGGTAATATCGTGTAGGTAGTATGATCGCTTGGGTCCTGTAAGAACCTCTCG 1144  
 QY 141 GluMetTyrAspValPheCysTyrArg-----Met 150  
 DB 1145 GAGCTCTGGAGTGCCTACTGTACCGCTGCAAGGTAGCTCCCTTCGTCGATCTCTGCC 1204  
 QY 150 -----Met 150  
 DB 1205 TACTCTCGGCCCTGAGCCCTTAATAAGCCTCTGTGAGCCACTGACCCGCTTTCCTG 1264  
 QY 151 LysAspValAsnCysThr\*\*LysValGlyTyrValGlyAspGlyPheSer---TyrSer 169  
 DB 1265 CA-GACGTGGCTTCCAGTGTGCGGCGGCTCTGCTGGTGTGACGGATCAGCAGCTGCAAC 1323  
 QY 170 GlyAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGluVal 189  
 DB 1324 GGGAACTGCTTGATGCTTGGCGCCACTGCTCCACTTCTCCACTTCTATGGGATGCTG 1383  
 QY 190 LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAspLeu 209  
 DB 1384 CTGGCTATGCAATGCCACCGACAGAGTCTTGAATTTATGATTTCTCTGGAGGATGAG 1443  
 QY 210 SerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGluAsnGlu----- 227  
 DB 1444 CTCACCTACAAAGACACTTTCGTTCTGTCACAAAGGCTTTGGACAAACATGGTACG 1503  
 QY 227 ----- 227  
 DB 1504 GCAGGGCTTGGCGAGTTGCTCTCTGCTGAGCTCGGCCCAACCCAGGCTTCTCTCTGTC 1563  
 QY 228 -----ThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePhe 245  
 DB 1564 TCATAGACGCTGAGTGGCCAGATCTAGAACTCCACGCTTCTAATGCCACCTTCTCAGT 1623  
 QY 246 TyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuIle 265  
 DB 1624 GTCAT---GCCAGCGGGGACATATGCTTCTCGCCCACTCAGGTCTTAGCTCTTTCATA 1680  
 QY 266 ThrAspArgGlnAspProLeuHis----- 273  
 DB 1681 AGTGACATGGCCCTTGACAA-CACTTCTCTGGTTCCTCTGTCGTGCTGCGCTCAACTGA 1739

QY 274 ---ProThrGluThrArg----- 278  
 DB 1740 AACCTTACTGACTCAGCCCTTGTCTACCGGCTGGGCTGACTGAGTTCCTCCATAGGCC 1799  
 QY 279 -----CysValAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIle 296  
 DB 1800 CCGGGGCGAGCTGTGGTTAGCCACGTCATCGTGGGACATCATGGCTTTCAACGGCATC 1859  
 QY 297 ThrHisValIleSerArg\*\*LeuLysAlaPro----- 307  
 DB 1860 ATTCACTGCTGGCCACCCCTGCTCATGCTCCACAGACTAGGCGAGTCTGGGATCT 1919  
 QY 308 ---ProAlaProValThrLeu\*\*HisThrGlyLeuGly\*\*GlyIlePhe\*\*\*\*\*Ile 326  
 DB 1920 GAGCTCCACCTGTGGCATTA-----AGCTTGGGGTGTGTGAACCTCTGGAACA 1970  
 QY 327 Ile-----LeuValThrGlyAlaValAlaLeuAlaLysTyrSerTyrPheArgIleAsn 344  
 DB 1971 CTGCTGGGGTGTGGCGGAGCTCTC-----TACCTGGCTGCGCCA 2012  
 QY 345 ArgLysThrIleGlyPhe 350  
 DB 2013 GGCACCACTCCAGTTTC 2030

RESULT 10  
 T47504  
 LOCUS  
 DEFINITION  
 T47504 459 bp mRNA linear EST 01-FEB-1995  
 yb14f01.r1 Stratagene placenta (#937225) Homo sapiens cDNA clone  
 IMAGE:71161 5', similar to similar to SP:A41735 A41735  
 HYALURONATE-BINDING PROTEIN TSG-6 PRECURSOR, mRNA sequence.

ACCESSION  
 VERSION  
 T47504.1 GI:649484  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 459)  
 AUTHORS  
 Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
 Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,  
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,  
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
 Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,  
 Trevisan,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.  
 and Marra,M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)  
 97044478  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 Other ESTs: yb14f01.s1  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 1402  
 High quality sequence stops: 308 Source: IMAGE Consortium, LLNL This  
 clone is available royalty-free through LLNL; contact the IMAGE  
 Consortium (info@image.llnl.gov) for further information.  
 Insert length: 1402 Std Error: 0.00  
 Seq primer: M13RP1  
 High quality sequence stop: 308.  
 Location/Qualifiers  
 1..459  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:492058"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:71161"  
 /sex="male"  
 /lab\_host="SOLR cells (kanamycin resistant)"

```

/clone_lib="Stratagene placenta (#937225)"
/note="Organ: placenta; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Caucasian. Average insert size: 1.2 kb; Uni-ZAP
XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTT 3'"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.34e-57 Length: 459
Score: 557.50 Matches: 128
Percent Similarity: 85.06% Conservative: 3
Best Local Similarity: 83.12% Mismatches: 16
Query Match: 29.89% Indels: 9
DB: 14 Gaps: 1

```

US-09-466-778B-11 (1-353) x T47504 (1-459)

```

Qy 88 TyrAsnGlnLeuSerTyr**GlnIysAlaLysTyrHisLeuCysSerAlaGlyTrieu 107
Db 2 TACAAACAGCTCTCTATGCCCAGAGGC-AAGTACCACCTGTCTCAGCAGGCTGGCTG 60
Qy 108 GlnThrGlyArgValAlaTyrProThrAlaPheLaSerGlnAsnCysGlySerGlyVal 127
Db 61 GAGACCGGGCGGGTTCCTACCCACAGCCTTCGCTCCAGAACTGTGGCTGTGGTGTG 120
Qy 128 ValGlyIleValAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCys 147
Db 121 GTTGGGATGATGGACTATGGACCTAGACCCACAGAGTGAATGTGGGATGTCTTCGTC 180
Qy 148 TyrArgMetLysAspValAsnCysThr**LysValGlyTyrValGlyAspGlyPheSer 167
Db 181 TATCGGATGAAGATGTGAACCTGACCTGCACCTG-CAAGTNGCTATGTGGGAGATGGCTTCTCA 239
Qy 168 TyrSer-GlyAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuTh 187
Db 240 TGCAGTGGGAACTGCTGCAGGTCCTGATGCTCTCTCCCTCACTCACAAATCTTCGTAC 299
Qy 197 rGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeu--GluHisLeu 206
Db 300 GGAAGTGTGGCTATTCCACAGCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 359
Qy 207 ThrAspLeu-SerIleArgGlyThrLeuPhe--ValProGlnAsnSerGlyLeuGlyCly 225
Db 360 ATTGACCTGTGTCATCCATCCCGGCGACCTTTTGTGTGCCACAGAACAGTNGGCTTGGGGG 419
Qy 226 Asn---GluThrLeuSerGlyArgAspIle 234
Db 420 AGATTGAGACNTTTTGTGGGGGGGACATC 449

```

## RESULT 11

## AK028441

## LOCUS

AK028441 3454 bp mRNA linear HTC 18-SEP-2003  
Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched  
library, clone:4631403B18 product:hypothetical Beta-Ig-H3/Fasciclin  
domain containing protein, full insert sequence.

## ACCESSION

## AK028441

## VERSION

## AK028441.1

## GI:26080850

## KEYWORDS

## HTC; CAP trapper.

## SOURCE

## Mus musculus

## ORGANISM

## Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## REFERENCE

## 1

## AUTHORS

## Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K.,

## Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

## JOURNAL

## Meth. Enzymol. 303, 19-44 (1999)

## MEDLINE

## 99279253

## PUBMED

## 10349636

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

## REFERENCE

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## FANTOM Consortium.

## Functional annotation of a full-length mouse cDNA collection

## Nature 409, 685-690 (2001)

## REFERENCE

## AUTHORS

## The RIKEN Consortium and the RIKEN Genome Exploration Research

## Group Phase I &amp; II Team.

## Analysis of the mouse transcriptome based on functional annotation

## of 60,770 full-length cDNAs

## Nature 420, 563-573 (2002)

## REFERENCE

## AUTHORS

## Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

## Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

## Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

## Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

## Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

## Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

## Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

## Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

## Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

## Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

## Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

## Muramatsu, M. and Hayashizaki, Y.

## Direct Submission

## Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

## Physical and Chemical Research (RIKEN), Laboratory for Genome

## Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

## RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

## cDNA library was prepared and sequenced in Mouse Genome

## Encyclopedia Project of Genome Exploration Research Group in Riken

## Genomic Sciences Center and Genome Science Laboratory in RIKEN.

## Division of Experimental Animal Research in Riken contributed to

## prepare mouse tissues.

## Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

## source

## 1. .3454

## /organism="Mus musculus"

## /mol\_type="mRNA"

## /strain="C57BL/6J"

## /db\_xref="FANTOM,DB:4631403B18"

## /db\_xref="MGI:2390349"

## /db\_xref="taxon:10090"

## /clone="4631403B18"

## /tissue\_type="skin"

## /clone\_lib="RIKEN full-length enriched mouse cDNA library"

## /dev\_stage="0 day neonate"

## misc\_feature

## 1. .3454

## /note="hypothetical Beta-Ig-H3/Fasciclin domain containing

## protein (InterPro|IPR000782, evidence: InterPro)"

## ORIGIN

## Normalization and subtraction of cap-trapper-selected cDNAs to

Alignment Scores:

Pred. No.:	2,028-55	Length:	3454
Score:	52.00	Matches:	131
Percent Similarity:	45.7%	Conservative:	48
Best Local Similarity:	33.5%	Mismatches:	126
Query Match:	29.6%	Indels:	88
DB:	11	Gaps:	7

US-09-466-778B-11 (1-353) x AK028441 (1-3454)

Qy	8	LysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsnCys---	GluproGluGln 26
Db	1756	CGCTGTGAATGCCACGTAGCGTACGTGGGTGATGGCTGCAGGTCTCAGAGGAGCTTTAAA	1815
Qy	27	LeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLysCysVal 46	
Db	1816	CCCCCTGTGACAGATGCTTTGGGAGGATCATCCACTTGGCCACACTGATGCTTTGTGCAC	1875
Qy	47	AspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeuGlyGln 66	
Db	1876	GACCTACATTTCCAGGAAAAACAGCGTGGTCTTCCACATCCAGGCCACACAGTGGCCCT	1935
Qy	67	TyrIysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAla 86	
Db	1936	TATGGCTGACCTTCTCAGAGCCACAGAGCGTGTGAGGCCAGGAGCGCTCTTGCT	1995
Qy	87	ThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyr 106	
Db	1996	TCACCTCCCTCAACTCTCTGCCGCCAACACAGCTAGGTTTTCATGTCTCTCTCGTGGGCTGG	2055
Qy	107	LeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGly 126	
Db	2056	TTGGCCAAACGGCTCTGCTGCCACCCTCTGTCACCCACGCGGACACTGTGGTAATAAT	2115
Qy	127	ValValGlyIleValAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPhe 146	
Db	2116	CGTGTAGGTGTAGTTAGCCTTGGGGTCCGTAAAGAACTCTCGAGACTCTGGGATGCCTAC	2175
Qy	147	CysTyrArg-----	149
Db	2176	TGTTACCGCGTGCAAGGTACGTCCCTTCCTCCGATCTCTGCCTACCTCCGCCCCCTGAG	2235
Qy	150	-----MetLysAspValAsnCysThr 156	
Db	2236	CCCCCTTAAGCCTCTGTGAGCCACTGACCCGGCTTTCCTGCA-GACGTGGCTTGCAG	2294
Qy	157	**LysValGlyTyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnVal 175	
Db	2295	TGTCGGCGCGCTTCGTGGGTGACGGGATCAGCACGTCGCAACGGGAAACTGCTTGATGTC	2354
Qy	176	LeuMetSerPheProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSer 195	
Db	2355	TTGGCCGCCACTGCAACTTCTCCACCTTCTATGGGATGTGCTGGGCTATGCCAATGCC	2414
Qy	196	SerAlaArgGlyArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeu 215	
Db	2415	ACCAGAGAGGTCTGTAATTATGGATTCTCTGGAGGACGAGCTCACTACAGACACTC	2474
Qy	216	PheValProGlnAsnSerGlyLeuGlyGluAsnGlu-----	227
Db	2475	TTCGTCTCTGTCAACAAAGGCTTTGTGGACAACATGGTTAAGCGGAGGCGCTTTGGCGAGTT	2534
Qy	228	-----ThrLeuSerGly 231	
Db	2535	TGCTCCCTGTGAGTGGGCCCCAACCCACACAGCTCTTGCTCTGCTCATAGACGCTGAGTGGC	2594
Qy	232	ArgAspIleGluHisHisLeuAlaAsnValSerMetPhePheTyrAsnAspLeuValAsn 251	
Db	2595	CCAGATCTAGAACTCCACGCTTCTTAACGCCACCTTCTGAGTATCAAT---GCCAGCGGG	2651
Qy	252	GlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuIleThrAspArgGlnAspPro 271	

```

Db      29  GATGACTGCAGTGTCTGGAGAGTGGGAACACACTGTGGACCGCTGCTGGGCCAGCCA  88
Qy      37  GlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGln----- 51
Db      89  CGCCCTCCCACTCAGATGCATGTGCACTGACCTGCATCCAGGGGTGTGCCCTTCG  148
Qy      51  ----- 51
Db      149  CCATCCCATGTPACCATTTTCATCTCAGAGCCGCGCAGCTGACCATGCACCCCTCCAT  208
Qy      52  -----AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeu  69
Db      209  TCTGCAGAAACGGCTGGCTTTTCCACCTCCAGGCCACCGCGGCCCTTATGGTCTG  268
Qy      70  ThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsn  89
Db      269  AACTTTTCGGAGCTGAGCGCATGCGA-ASACAGGAGCGCTGCTTCTTCATTCCT  327
Qy      90  GlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThr  109
Db      328  CAGCTCTCTGCTGCCAGCAGCTGGCTTCCACCTGTGCCTCATGGCTGGCTGCCAAT  387
Qy      110  GlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGly  129
Db      388  GGCTCCACTGCCACCTGTGTTTCCCTGTGGCGGACTGTGGCAATGGTGGGTGGGC  447
Qy      130  IleValAspTyrGlyProArgProLysSerGluMetTrpAspValPheCysTyrArg  149
Db      448  GTAGTCAGCTGGTGGTCCCGCAAGAACCTCTCAGAACGCTGGGATGCCCTACTGCTTCGT  507
Qy      150  MetLysAspValAsnCysThr***LysValGlyTyrValGlyAspGlyPheSer---Tyr  168
Db      508  GTGCAAGATGTGGCTGCCGATGCCGAATGCTTCGTGGGTGACGGATCAGCAGTGC  567
Qy      169  SerGlyAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGlu  188
Db      568  AATGGGAAGCTGCTGGATGTGCTGGCTGCCACTGCCAACTTCTCCACCTTCTATGGATG  627
Qy      189  ValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAsp  208
Db      628  CTATTGGGTATGCCAATGCCACCCAGCGGGTCTCGACTTCCCTGCACTTCTGATGAT  687
Qy      209  LeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGluAsnGluThr  228
Db      688  GAGCTCACATATAAGACATCTTCGTCCCTGTCAATGAAGGCTTTGTGGACAACATGAGC  747
Qy      229  LeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPheTyrAsnAsp  248
Db      748  CTGAGTGGCCAGACTTGGAGCTGCATGCTCCACGCCACCTCCTAAGTGCCCAAC--- 804
Qy      249  LeuValAsnGlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuIleThrAspArg  268
Db      805  GCACGACGAGGGAAGTTGCTTCGGGCCACTCAGGCCTCAGGCTCATCATCATGACGCA  864
Qy      269  GlnAspProLeuHisProThrGluThrArgCysValAspGlyArgAspThrLeuGluTrp  288
Db      865  -----GGCCCTGACACAGTTCCTGG  885

RESULT 13
CB268774
LOCUS
DEFINITION
1007680 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
CDNA 5', mRNA sequence.
CB268774
ACCESSION
CB268774.1 GI:28443359
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 572)
AUTHORS
Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE
EST analysis of human adipose gene expression

```

```

JOURNAL COMMENT
Unpublished (2002)
Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTGGGAAGCGCCCATTTGTGTGGT
BACKWARD: AATCACTACTATAGGCGCAATTGG
Seq primer: GTTGTACCGCGGAATTC.
FEATURES
Location/Qualifiers
1..572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambdaTriplex"

ORIGIN
Alignment Scores:
Pred. No.: 1,04e-45 Length: 572
Score: 462.00 Matches: 86
Percent Similarity: 62.57% Conservative: 31
Best local Similarity: 45.99% Mismatches: 68
Query Match: 24.77% Indels: 2
DB: 14 Gaps: 2

US-09-466-778B-11 (1-353) x CB268774 (1-572)
Qy 15 TyrValGlyAspGlyLeuAsnCys---GluProGluGlnLeuProIleAspArgCysLeu 33
Db 6 TACGTAGCGATGGACTGCTGAGGAGTGGGAACCCACTGTGGACCGCTGCTTG 65
Qy 34 GlnAspAenGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGlnAspThr 53
Db 66 GGCAGCCACCGCCCTGCCACTCAGATGCCATGTGCATGACCTGCATTCAGGAGAA 125
Qy 54 ThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPheAspLys 73
Db 126 CGGCTGGCGTTTCCACCTCCAGGCCACCGGCCCTTATGGTCTGAACCTTTTCGGAG 185
Qy 74 AlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGlnLeuSerTyr 93
Db 186 GCTGAGCGCGCATGCGAAGCAGCAGGAGCCCTGCTTCTTCATTCCTCCCTCAGCTCTCTGCT 245
Qy 94 ***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGlyArgValAla 113
Db 246 GCCCAGCAGCTGGGCTTCCACCTGTGCTCATGGCTGGCTGGCCATGGCTCCACTGCC 305
Qy 114 TyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIleValAspTyr 133
Db 306 CACCTGTGGTTTCCCTGTGGCGACGTGGCAATGTCTGGTGGGGGTAGTCAGCCTG 365
Qy 134 GlyProArgProAsnLysSerGluMetTrpAspValPheCysTyrArgMetLysAspVal 153
Db 366 GGTGCCCGCAAGAACCTCTCAGAACGCTGGATGCTACTGCTTCCGTGTCAGAGATGG 425
Qy 154 AsnCysThr***LysValGlyTyrValGlyAspGlyPheSer---TyrSerGlyAsnLeu 172
Db 426 GCCTGCCATGCCGAATGGCTTGTGGTACGGGATCAGCAGCTGCAGTGGGAGCTG 485
Qy 173 LeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGluValLeuAlaTyr 192
Db 486 CTGGATGTGCTGGCTGCCACTGCCAATCTCTCCACCTTCTATGGATGCTATGGCTAT 545
Qy 193 SerAsnSerSerAlaArgGly 199
Db 546 GCCAATGCCACCCACGGGT 566

```

```

RESULT 14
BX705197      865 bp      mRNA      linear      EST 17-NOV-2003
LOCUS      BX705197
DEFINITION      BX705197 XGC-tadpole Silurana tropicalis cDNA clone TTPA006003 5',
                mRNA sequence.
ACCESSION      BX705197
VERSION      BX705197.1 GI:38367404
KEYWORDS      EST.
SOURCE      Silurana tropicalis (western clawed frog)
ORGANISM      Silurana tropicalis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
              Xenopodinae; Silurana.
REFERENCE      1 (bases 1 to 865)
AUTHORS      Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE      Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL      Unpublished (2003)
COMMENT      Contact: Croning MDR
              Sanger Institute
              Hinxton, Cambridgeshire, CB10 1SA, UK
              Email: trop@sanger.ac.uk
              Sanger Xenopus tropicalis EST project 2001
              TROPICALIS_SEQUENCE ID: TTPA006003.p1kSP6
              Sequencing primer: SP6
              This sequence is from a Xenopus Gene Collection (XGC) library
              constructed by Nigel Garrett.
              cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole
              embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
              EcoRI at the 5' end and NotI at the 3' end.
              Vector: pCS107; Site1: EcoRI; Site2: NotI.
              Host: Escherichia coli DH10B.
              Location/Qualifiers
                1..865
                /organism="Silurana tropicalis"
                /mol_type="mRNA"
                /db_xref="taxon:8364"
                /clone="TTPA006003"
                /dev_stage="tadpole (stage 35-40)"
                /lab_host="E. coli DH10B"
                /clone_lib="XGC-tadpole"
                /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
                was oligo dt primed from 5ug of poly A+ RNA from tadpole
                embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
                with EcoRI at the 5' end and NotI at the 3' end"

FEATURES
    source
    1..865
    /organism="Silurana tropicalis"
    /mol_type="mRNA"
    /db_xref="taxon:8364"
    /clone="TTPA006003"
    /dev_stage="tadpole (stage 35-40)"
    /lab_host="E. coli DH10B"
    /clone_lib="XGC-tadpole"
    /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
    was oligo dt primed from 5ug of poly A+ RNA from tadpole
    embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
    with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Alignment Scores:
Pred. No.:      6.96e-43      Length:      865
Score:          441.50      Matches:      89
Percent Similarity: 68.15%      Conservative: 18
Best Local Similarity: 56.69%      Mismatches: 43
Query Match:    23.67%      Indels:      7
DB:             13      Gaps:      2

US-09-466-778B-11 (1-353) x BX705197 (1-865)

Qy 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219
Db 1 AAAGAAATTTTCACTACCTACTCACTAATCTATCAGTGCAAGCGACTCTGTTGGCCCAAGC 60
Qy 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
Db 61 AATGATGGACTTAATGAATATCAGACCCCTGCTGGACCGCATATAGAAATATCATCTGGCA 120
Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
Db 121 AATGTCAGCAAGTTCTTCTTGAGCAAGCTTGACCAATGGACACTCTTCACACCAAGATA 180
Qy 260 GlySerLysLeuLeuIleThrAspArgGlnAppProLeuHisPro----- 274
Db 181 GGCCCAAGTTATTAATTTCTTTGACAAATGACCCAGCCAGCAAAATAACCACTGATTCC 240
Qy 275 ---ThrGluThrArgCysValAspGlyArgAspThrLeuGluTrpAspIleCysAlaSer 293

```

```

Db 241 ATGACTCCAAAGATATGTGGATGGAAACCTATTTTGCAGTGGGACATATTCGCTTCC 300
Qy 294 AsnGlyIleThrHisValIleSerArg***LeuLysAlaProAlaProAlaProValThrLeu 313
Db 301 AATGGATCATTCACACAATTTGCTGAACCTCTTAACCTCGGCTCCAGAGCAATAGCTCTG 360
Qy 314 ***HisThrGlyLeuGly***GlyIlePhe*****IleIleLeuValThrGlyAlaVal 333
Db 361 ---CATGCTGGCCATCGAGCAGGATATTTCTTTGGCAATTTGTTGATCGTCGGGCTCCTG 417
Qy 334 AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgLysThrIleGlyPhe 350
Db 418 GCATCGCAGTACTATTATTAAGAAGTTCACAGAAAGATTTCCAATTC 468

CA342417      719 bp      mRNA      linear      EST 05-NOV-2002
LOCUS      CA342417
DEFINITION      672284 NCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT32P18_D_H09 5',
                mRNA sequence.
ACCESSION      CA342417
VERSION      CA342417.1 GI:24587588
KEYWORDS      EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
ORGANISM      Oncorhynchus mykiss
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei;
              Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
              1 (bases 1 to 719)
              Rexroad,C.E. and Keele,J.W.
              Sequence analysis of a rainbow trout normalized cDNA library
              Unpublished (2002)
              Contact: Rexroad CE
              USDA, ARS, National Center for Cool and Cold Water Aquaculture
              11876 Lestown Road, Kearneysville, WV 25430, USA
              Tel: 304 724 8340 x2129
              Fax: 304 725 0351
              Email: crexroad@nccwa.ars.usda.gov
              Single pass sequencing. Bases called with phred v0.020425.c and
              trimmed with the aid of the trim_alt option. Vector identified by
              cross match v0.990329.
              Seq primer: AGCGATAACAATTTTCACACAGGA.
              Location/Qualifiers
                1..719
                /organism="Oncorhynchus mykiss"
                /mol_type="mRNA"
                /db_xref="taxon:8022"
                /clone="1RT32P18_D_H09"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /clone_lib="NCCWA 1RT"
                /note="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
                Library made from pooled tissue from brain, gill, liver,
                spleen, muscle, and kidney."

ORIGIN
Alignment Scores:
Pred. No.:      5.84e-43      Length:      719
Score:          441.00      Matches:      96
Percent Similarity: 56.90%      Conservative: 40
Best Local Similarity: 40.17%      Mismatches: 97
Query Match:    23.65%      Indels:      6
DB:             14      Gaps:      2

US-09-466-778B-11 (1-353) x CA342417 (1-719)

Qy 67 TyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAla 86
Db 2 TACCAGCTGAAGTAACCGAGCTGAGCTAGCTAGCTGCTGAGCTGAGGGGCTACGCTGCC 61
Qy 87 ThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrp 106
Db 62 ACCTTCTCAGATGGCTGATGTCACACAGCTGGGAATGCACCGATGTGTGCTGGATGG 121

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 1, 2004, 11:22:04 ; Search time 1504 Seconds  
(without alignments)  
877.740 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPGHKCEKSHYVDGL.....ALAAYSYFRINRKTIGFHF 353

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2465228 seqs, 1869859620 residues

Total number of hits satisfying chosen parameters: 4930456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09466778/runat\_01042004\_095644\_10700/app.query.fasta\_1.519  
-DB=Published Applications NA -CFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pcr -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODEB=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09466778 @CEN 1 1 396 @runat\_01042004\_095644\_10700  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubnpa/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubnpa/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubnpa/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/2/pubnpa/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubnpa/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/2/pubnpa/US09C\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/2/pubnpa/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/2/pubnpa/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/2/pubnpa/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/2/pubnpa/US10C\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq\*
- 18: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query  
No. Score Match Length DB ID Description

1	1772	95.0	4576	9	US-09-842-930A-24	Sequence 24, Appl
2	1772	95.0	4642	14	US-10-133-172-3	Sequence 3, Appl
3	1772	95.0	4962	14	US-10-133-172-19	Sequence 19, Appl
4	1769	94.9	3260	15	US-10-028-248A-210	Sequence 210, Appl
5	1769	94.9	3260	15	US-10-107-782-210	Sequence 210, Appl
6	1558	83.5	8495	15	US-10-028-248A-3	Sequence 3, Appl
7	1558	83.5	8495	15	US-10-107-782-3	Sequence 3, Appl
8	1554	83.3	8444	15	US-10-028-248A-1	Sequence 1, Appl
9	1554	83.3	8444	15	US-10-107-782-1	Sequence 1, Appl
10	1514.5	81.2	4706	9	US-09-842-930A-1	Sequence 1, Appl
11	1514.5	81.2	4706	14	US-10-133-172-1	Sequence 1, Appl
12	661.5	35.5	3681	15	US-10-264-237-1320	Sequence 1320, Ap
13	645.5	34.6	2753	10	US-09-774-639-13	Sequence 12, Appl
14	645.5	34.6	2753	10	US-09-969-730-13	Sequence 13, Appl
15	645.5	34.6	2753	16	US-10-621-363-13	Sequence 1824, Ap
16	642	34.4	2160	15	US-10-104-047-1824	Sequence 1824, Ap
17	632.5	33.9	1482	15	US-10-305-720-230	Sequence 230, Appl
18	581	31.2	2883	15	US-10-104-047-610	Sequence 610, Appl
19	363	19.5	474	10	US-09-918-995-32777	Sequence 32777, A
20	267.5	14.3	1144	13	US-10-044-090-446	Sequence 446, Appl
21	267.5	14.3	1144	14	US-10-247-671-124	Sequence 124, Appl
22	267.5	14.3	1422	15	US-10-295-027-291	Sequence 291, Appl
23	267.5	14.3	1430	15	US-10-295-027-1020	Sequence 1020, Ap
24	267.5	14.3	1728	13	US-10-044-090-445	Sequence 445, Appl
25	266.5	14.3	1414	9	US-09-799-118-1	Sequence 1, Appl
26	262	14.0	466	12	US-10-085-783A-58386	Sequence 58386, A
27	262	14.0	466	15	US-10-242-535A-58386	Sequence 58386, A
28	256.5	13.8	1411	9	US-09-765-231A-44	Sequence 44, Appl
29	248	13.3	351	14	US-10-060-036-880	Sequence 880, Appl
30	243.5	13.1	1414	15	US-10-133-937-18	Sequence 18, Appl
31	243.5	13.1	1414	15	US-10-295-027-293	Sequence 293, Appl
32	243.5	13.1	1414	15	US-10-159-563-18	Sequence 18, Appl
33	231	12.4	396	10	US-09-918-995-17767	Sequence 17767, A
34	213	11.4	718	14	US-10-023-896-38	Sequence 38, Appl
35	202.5	10.9	1365	13	US-10-098-841-328	Sequence 328, Appl
36	200.5	10.8	1985	9	US-09-909-320-212	Sequence 212, Appl
37	200.5	10.8	1985	9	US-09-909-088B-212	Sequence 212, Appl
38	200.5	10.8	1985	9	US-09-905-291A-212	Sequence 212, Appl
39	200.5	10.8	1985	9	US-09-902-853-212	Sequence 212, Appl
40	200.5	10.8	1985	9	US-09-907-824-212	Sequence 212, Appl
41	200.5	10.8	1985	9	US-09-907-841-212	Sequence 212, Appl
42	200.5	10.8	1985	10	US-09-904-011-212	Sequence 212, Appl
43	200.5	10.8	1985	10	US-09-906-742-212	Sequence 212, Appl
44	200.5	10.8	1985	10	US-09-906-838-212	Sequence 212, Appl
45	200.5	10.8	1985	10	US-09-907-613-212	Sequence 212, Appl

ALIGNMENTS

RESULT 1  
US-09-842-930A-24  
; Sequence 24, Application US/09842930A  
; Publication No. US20020197681A1  
; GENERAL INFORMATION:  
; APPLICANT: Weigel, Paul  
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis  
; FILE REFERENCE: 5820.603  
; CURRENT APPLICATION NUMBER: US/09/842.930A  
; CURRENT FILING DATE: 2001-04-22  
; PRIOR APPLICATION NUMBER: 60/245,320  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: 60/199,538  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 4576  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-842-930A-24

Alignment Scores: 3.72e-231 Length: 4576  
Pred. No.: 3.72e-231 Length: 4576

```
Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservative: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x US-09-842-930A-24 (1-4576)

Qy 1 MetThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 2956 ATGACAGCCCGGGCAACCAAGTGTGAGTGAAGTCACTATGTCGGAGATGGGCTG 3015

Qy 21 AsnCysGluProGluLeuProLeuAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 3016 AACTGTGACCGGAGCAGCTGCCATTCACCGCTGCTTACAGGACAATGGCAGTGCAT 3075

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 3076 GCAGAGCCCAATGTGTCGACCTCCACTCCAGGATACACTGTGGGTGTTCACTA 3135

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 3136 CGCTCCCACTGGGCGCAGTATAAGCTGACCTTGACAAAGCCAGAGGCGCTGTGCCAAC 3195

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 3196 GAAGCTGGACCATGGCAACCTACAAACAGCTCTCTATGCCAGAGGCCAAGTACCCAC 3255

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 3256 CTGTGCTCAGCAGGCTGCTGGAGCCGGGGGTGCTACCCACAGCCCTTGCCTCC 3315

Qy 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
Db 3316 CAGAACTGTGCTGTGTTGGGTAGTGGACTATGACCTATGACCTAGACCCCAACAAGAGT 3375

Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 3376 GAAATGTGGAGTGTCTCTGCTATCGGATGAAGATGAACTGCACCTGCAAGGTGGC 3435

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 3436 TATGTGGAGATGCTTCTCATGAGTGGGAACCTGCTGAGGCTCTCATGCTCTCC 3495

Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 3496 TCACCTCAAACTTCCTGACGGAAGTGTGGCTATTCACACAGCTCAGCTCGAGGCGGT 3555

Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 3556 GCATTTCTAGAACACCTGACCTGCTCCATCCGCGGACCCCTTTGTGGCCACAGAAC 3615

Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240
Db 3616 AGTGGCTGGGGAGATGACCTGTGCTGGCGGACATCGAGCACCACTCGCCAAAT 3675

Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 3676 GTCAGCATGTTTCTACAAATGACCTGTGCAATGGCACCACTCCGCAACAGAGCTGGGA 3735

Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 3736 AGCAAGCTGCTCATCTGCGCAGCAGGACCCACTCCAAACGAGGACCAAGTTGTT 3795

Qy 281 AspGlyArgAspThrLeuLeuTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 3796 GATGGAAGAGCATTCTTCAGTGGGACATCTTGGCTCCATGGATCATTCATGTCATT 3855

Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 3856 TCCAGGCGCTTTAAAGAGCACCCCTGCCCCGCTGACCTTGACCCCACTGGCTGGGAGCA 3915

Qy 321 GlyIlePhe*****IleIleValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
Db 3915
```

```
Db 3916 GGGATCTTCTTGGCATCATCTGGTGAAGTGGGCTGTTCCTGGCTGCTTACTCTAC 3975

Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 3976 TTTCCGATTAACCGGAGAACAAATCGGCTTCCAGCATTTT 4014

RESULT 2
US-10-133-172-3
; Sequence 3, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-172-3

Alignment Scores:
Pred. No.: 3,81e-231 Length: 4642
Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservative: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 14 Gaps: 0

US-09-466-778B-11 (1-353) x US-10-133-172-3 (1-4642)

Qy 1 MetThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 3022 ATGACAGCCCGGGCAACCAAGTGTGAGTGAAGTCACTATGTCGGAGATGGGCTG 3081

Qy 21 AsnCysGluProGluLeuProLeuAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 3082 AACTGTGACCGGAGCAGCTGCCATTCACCGCTGCTTACAGGACAATGGCAGTGCAT 3141

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 3142 GCAGAGCCCAATGTGTCGACCTCCACTCCAGATACCACTGTGGGGTGTTCATCTA 3201

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 3202 CGCTCCCACTGGGCGCAGTATAAGTGCACCTTTGACAAAGCCAGAGGCGCTGTGCCAAC 3261

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 3262 GAAGCTCGCAGCATGGCAACCTACCAACAGCTCTCTATGCCAGAGGCGCAAGTACCCAC 3321

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 3322 CTGTGCTCAGCAGGCTGCTGGAGACCGGGCGGTGCTTACCCACAGCTTTCGCTCC 3381

Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 3382 CAGACTGTGCTGCTGTGTTGGATAGTGGACTATGACCTAGACCTACACCAACAGAGT 3441

Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 3442 GAAATGTGGAGTGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCACCTGCAAGGTGGC 3501

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 3501
```

Db 3502 TATGTGGAGATGGCTTCTCATGAGTGGGAGCACTGCTGAGGTCCTGAGTCTCTCC 3561  
 Qy 181 SerLeuThrAnPheLeuThrGluValLeuAlaTyrSerAnSerAlaAtrGlyArg 200  
 Db 3562 TCACTCAAACTTCTTGACGGAAGTGTGCTTATTCACAGCTCAGCTCGAGGCGGT 3621  
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 Db 3622 GCATTTCTAGAACACTGACTGACCTGTCCATCCGCGCACCTCTTTGTGCCACAGAAC 3681  
 Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 Db 3682 AGTGGGCTGGGGAGATGAGACCTTGTCTGGGGGAGACATCGAGCACCACTCGCCAAT 3741  
 Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 Db 3742 GTACGATGTTTCTACATGACCTTGTCAATGGCACCCCTGCAACGAGGCTGGGA 3801  
 Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 Db 3802 AGCAAGCTGCTCATCACTGCGCAGCAGGACCACTCCCAACCGAGAGACCAAGTTGTT 3861  
 Qy 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValle 300  
 Db 3862 GATGGAGAGACCATCTCTGCACTGGGAGCATCTTTGCCCTCCAAATGGGATCATTCATGTCT 3921  
 Qy 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 3922 TCCAGGCTTTAAAGACACCCCTGCCCCCTGACCTTGACCCACACTGGCTTGGGAGCA 3981  
 Qy 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
 Db 3982 GGGATCTTCTTTGCCATCATCTGCTGACTGGGCTGTGCTTGGCTTGGCTTACTCTCTAC 4041  
 Qy 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 4042 TTTCCGATAAACCGGAGAACAACTCGGCTTCCAGCATTTT 4080

RESULT 3

US-10-133-172-19  
 ; Sequence 19, Application US/10133172  
 ; Publication No. US20030104987A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WEIGEL, PAUL H  
 ; APPLICANT: WEIGEL, JANET A  
 ; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS  
 ; FILE REFERENCE: 5864.014  
 ; CURRENT APPLICATION NUMBER: US/10/133,172  
 ; PRIOR FILING DATE: 2002-04-25  
 ; PRIOR APPLICATION NUMBER: 60/286,468  
 ; PRIOR FILING DATE: 2001-04-25  
 ; PRIOR APPLICATION NUMBER: 09/842,930  
 ; PRIOR FILING DATE: 2001-04-25  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 19  
 ; LENGTH: 4962  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-133-172-19

Alignment Scores:

Pred. No.: 4.24e-231 Length: 4962  
 Score: 1772.00 Matches: 334  
 Percent Similarity: 95.18% Conservative: 2  
 Best Local Similarity: 94.62% Mismatches: 17  
 Query Match: 95.01% Indels: 0  
 DB: 14 Gaps: 0

US-09-466-778B-11 (1-353) x US-10-133-172-19 (1-4962)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20

Db 3733 ATGACAGGCCCGCGCAAGCAAGTGTGAGTGTAAAGTCACTATGTTCGAGATGGCGTG 3792  
 Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 Db 3793 AACTGTGAGCGGAGAGAGCTGCCANTGACCGCTGTACAGGACAAATGGCGAGTGCAT 3852  
 Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 3853 GCAGAGCCCAATGTGTGCGACCTCCACTCCAGGATACCACTGTGGGGTGTTCATCTA 3912  
 Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 3913 CGTCCCCACTGGCCAGTATTAAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 3972  
 Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 Db 3973 GAAGCTGGACCAATGGCACTTACACCACTCTCTATGCCAGAGGCAAGTACCACTAC 4032  
 Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 4033 CTGTGCTCAGAGGCTGGCTGGAGACCGGGCGGGTGGCTTACCCACAGCCTTCGCCCTC 4092  
 Qy 121 GluAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 4093 CAGAACTGTGGCTCTGTGTGGTGGATAGTGGACTATGGACTAGACCCCAACAAAGAGT 4152  
 Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 Db 4153 GAAATGTGGAGTGTCTTCTGTATCGATGAAAGATGTGAATGCACTGCAGGTGGC 4212  
 Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180  
 Db 4213 TATGTGGAGATGGCTTCTCATGAGTGGAACTGCTGAGGTGCTGTGTCTTCTTCCCC 4272  
 Qy 181 SerLeuThrAnPheLeuThrGluValLeuAlaTyrSerAnSerSerAlaAtrGlyArg 200  
 Db 4273 TCACCTCAAACTTCTCGAGGAGTGTGCTGCTATTCCAAACAGCTCAGCTCGAGGCGGT 4332  
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 Db 4333 GCATTTCTAGAACACTGACTGACCTGCTCATCCGCGGCACTCTTTGTGCCACAGAAC 4392  
 Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 Db 4393 AGTGGGCTGGGGAGATGAGACCTTGTCTGGGGGAGACATCGAGCACCACTCGCCAAT 4452  
 Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 Db 4453 GTACGATGTTTCTACATGACCTTGTCAATGGCACCCCTGCAACCGAGGAGGCTGGGA 4512  
 Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 Db 4513 AGCAAGCTGCTCATCACTGCGCAGCAGGACCCACTCCCAACCGAGGAGACCAAGTTGTT 4572  
 Qy 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValle 300  
 Db 4573 GATGGAGAGACCATCTGCACTGGGAGCATCTTTGCCCTCCAATGGGATCATTCATGTCT 4632  
 Qy 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 4633 TCCAGGCTTTAAAGACACCCCTGCCCCCTGACCTTGACCTTGACCCACACTGGCTGGGAGCA 4692  
 Qy 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
 Db 4693 GGGATCTTCTTTGCCATCATCTGCTGAGTGGGGCTGTGCTTGGCTTGGCTTACTCTCTAC 4752  
 Qy 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 4753 TTTCCGATAAACCGGAGAACAACTCGGCTTCCAGCATTTT 4791

RESULT 4

US-10-028-248A-210

```

; Sequence 210, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkete, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytke, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennada
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: NO. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 210
; LENGTH: 3260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3171)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-210

Alignment Scores:
Pred. No.: 5,58e-231 Length: 3260
Score: 1769.00 Matches: 333
Percent Similarity: 95.18% Conservative: 3
Best Local Similarity: 94.33% Mismatches: 17
Query Match: 94.85% Indels: 0
DB: 15 Gaps: 0

US-09-466-778B-11 (1-353) x US-10-028-248A-210 (1-3260)
QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisThrValGlyAspGlyLeu 20
DB 1635 ATGACAGCGCCGGGCAAGCACAAAGTGTGAGTGTAAAGTCACTATGTCCGAGATGGGCTG 1694
QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 1695 AACTGTGACCGGAGAGAGTGGCCATTGACCGCTGTACAGACAAATGGGAGTGGCCAT 1754

```

```

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 1755 GCAGAGCCCAATGTGTGACCTCCACTTCAGGATACCACTGTGGGTGTTCATCTA 1814
QY 61 ArgSerProLeuGlyGlnThrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 1815 CGTCCCTCCACTGGGCCAGTATAGCTTGTGACAAAGCCAGAGAGGCTGTGCCAAC 1874
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 1875 GAAGCTCGGACCATGGCACTCAACACAGCTCTCTATGCCAGAGAGGCCAAGTACCAC 1934
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 1935 CTGTGCTCAGCAGGCTGGCTGGAGACGGGGTTCCTACCCACAGGCTTCGGCTCC 1994
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 1995 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGACCTAGACCCACCAAGAGT 2054
QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
DB 2055 GAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAATGCACTGCAAGTGGGC 2114
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
DB 2115 TATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGTGCAGGTCCTGATGTCTTCCCC 2174
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
DB 2175 TCACCTCAAACTTCTTCAGGGAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGCGGT 2234
QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
DB 2235 GCATTTCTAGAACACCTGACTGACCTGTCCATCGCGGCACCCCTCTTTGTGCCACAGAAC 2294
QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
DB 2295 AGTGGGCTGGGGAGAAATGAGACCTGTCTGGGGGGACATCGAGACCACTCGCCAAT 2354
QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
DB 2355 GTGAGCATGTTTTCTACATGACCTTGTCAATGGCACCCCTTGCACAAAGAGGGTGGGA 2414
QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
DB 2415 AGCAAGCTGTCTCATCTGCGCAGCCAGGACCCACTCCAAACCGAGAGACCAAGTGTGT 2474
QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
DB 2475 GATGGAAGAGCCATTTGCGATGGGAGACATTTGCCCTCCATGGGATCATTCATGTCAAT 2534
QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
DB 2535 TCCAGGCGCTTTAAAGCACCCCTGCCCCCTGACCTTGACCCACACTGGCTTGGGAGCA 2594
QY 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
DB 2595 GGGATCTTCTTTGGCATCATCTGTGACTGGGGCTGTGTGCTTGGTGTCTTACTCTCTAC 2654
QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
DB 2655 TTTTCGATAAACCGGAGAACCAATCGGCTTCCAGCATTTT 2693

```

```

RESULT 5
US-10-107-782-210
; Sequence 210, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven

```

APPLICANT: Edinger, Shlomit,  
APPLICANT: Gangolli, Baha,  
APPLICANT: Kekuda, Ramesh,  
APPLICANT: Li, Li,  
APPLICANT: Liu, Xiaohong,  
APPLICANT: Malyankar, Uriel,  
APPLICANT: Miller, Charles,  
APPLICANT: Millet, Isabelle,  
APPLICANT: Patturajan, Meera,  
APPLICANT: Rothenberg, Mark,  
APPLICANT: Sciore, Paul,  
APPLICANT: Shenoy, Suresh,  
APPLICANT: Shimkets, Richard,  
APPLICANT: Si, Jingsheng,  
APPLICANT: Smithson, Glennda,  
APPLICANT: Spytek, Kimberly,  
APPLICANT: Stone, David,  
APPLICANT: Taupier, Raymond, Jr.,  
APPLICANT: Tchernev, Velizar,  
APPLICANT: Vernet, Corine,  
APPLICANT: Zerhusen, Brian  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF  
FILE INVENTION: 21402-222CIP  
CURRENT APPLICATION NUMBER: US/10/107,782  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: 10/028,248  
PRIOR FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/256,619  
PRIOR FILING DATE: 2000-12-19  
PRIOR APPLICATION NUMBER: 60/262,959  
PRIOR FILING DATE: 2001-01-19  
PRIOR APPLICATION NUMBER: 60/272,408  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/285,189  
PRIOR FILING DATE: 2001-04-20  
PRIOR APPLICATION NUMBER: 60/308,039  
PRIOR FILING DATE: 2001-07-26  
PRIOR APPLICATION NUMBER: 60/311,266  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 60/279,344  
PRIOR FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 215  
SOFTWARE: Curation version 0.1  
SEQ ID NO 210  
LENGTH: 3260  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (171)..(2861)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (3171)..(3171)  
OTHER INFORMATION: Wherein n may be a, c, g or t  
US-10-107-782-210

Alignment Scores:  
Pred. No.: 5.58e-231 Length: 3260  
Score: 1769.00 Matches: 333  
Percent Similarity: 95.18% Conservative: 3  
Best Local Similarity: 94.33% Mismatches: 17  
Query Match: 94.85% Indels: 0  
DB: 15 Gaps: 0

US-09-466-778b-11 (1-353) x US-10-107-782-210 (1-3260)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
Db 1635 ATGACAGCCCGGCGAGCAACAGTGTGAGTGAATGCTACTATGCGGAGATGGGCTG 1694  
Qy 21 AsnCysGluProGluGlnLeuProileAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
Db 1695 AACTGTGAGCGGAGCAGCTGCCCATTTGACCGCTTACAGGACATGGCAGTGGCAT 1754

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGluAspThrThrValGlyValPheHisLeu 60  
Db 1755 GCAGAGCCCAATGTGTGACCTCCACTTCCAGATACCACTGTGGGGTGTTCATCTA 1814  
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
Db 1815 CGCTCCCACTGGCGCAGTATAAGCTGACCTTTGCAAGCCAGAGAGGCTGTGCCAAC 1874  
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
Db 1875 GAAGCTGGCAGCATGGCAACCTACACAGCTCTCTATGCCAGAGGCGCAAGTACCAC 1934  
Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaLysTyrProThrAlaPheAlaSer 120  
Db 1935 CTGTGCTCAGCAGCTGGCTGGAGACCGGGCGGGTTCCTTACCCCAAGCTTCGCTCC 1994  
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
Db 1995 CAGAACTGTGGCTCTGTGTGGTGGATAGTAGCTATGAGCTAGACCCCAAGAGT 2054  
Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
Db 2055 GAAATGTGGGATGCTTCTGCTATCGGATGAAAGATGTGAAGTGCACCTGCAAGGTGGC 2114  
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180  
Db 2115 TATGTGGGAGATGGCTTCTCATCAGTGGGAACCTGTGTCAGGTCTCTGTATGCTTCCC 2174  
Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
Db 2175 TCACCTCACAACTTCTGACGGAAGTGTGCTGCTATTCACAGCTCAGCTCAGGCGCT 2234  
Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
Db 2235 GCATTTCTAGACACACCTGACTGACCTGTCCATCGCGGCCACCTCTTTGTGCCACAGA 2294  
Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLysLeuAlaAsn 240  
Db 2295 AGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCACCTCGCCA 2354  
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
Db 2355 GTCAGCATGTTTTTCTACAAATGACCTTGTCAATGGCACCACTCCAAACAGAGGGTGGGA 2414  
Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
Db 2415 AGCAAGCTGCTCATCTGTCAGCAGGAGCCCACTCCAAACCGAGGAGACAGGTTGTT 2474  
Qy 281 AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
Db 2475 GATGGAAGAGCCATTTGCAAGTGGGACATCTTTGCCCTCCAATGGGATCATTCATGCTAT 2534  
Qy 301 SerArg\*\*\*LeuLysAlaProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
Db 2535 TCCAGGCTTTAAAGACACCCCTGCCCCGTGACCTTGACCCACACACTGGCTGGGAGCA 2594  
Qy 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaLysTyr 340  
Db 2595 GGGATCTCTTTGCCATCATCTCTGGTGTGCTGGGCTGTTCCTTGGCTGTGCTTACTCTAC 2654  
Qy 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
Db 2655 TTTTCGATAAACCGAGACAATCGCTTCCAGCATTTT 2693

## RESULT 6

US-10-028-248A-3  
; Sequence 3, Application US/10028248A  
; Publication No. US20030235892A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Vernet, Corine

```
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytke, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Coleman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 8495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131)..(1316)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-3

Alignment Scores:
Pred. No.: 2,13e-201 Length: 8495
Score: 1558.00 Matches: 300
Percent Similarity: 85.31% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 15
Query Match: 82.54% Indels: 37
DB: 15 Gaps: 1

US-09-466-778B-11 (1-353) x US-10-028-248A-3 (1-8495)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
DB 6948 ATGACAGGCCCCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGGTG 7007

QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAsnGlyGlnCysHis 40
DB 7008 AACTGTGAGCGGAGCAGCTGCCATTCAGCCCTGCTTACAGACAAATGGGCAAGTGCAT 7067

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 7068 GCAGACGCCAAATGTGTGACCTCCACTTCCAGGATACCACTGTGTGGGTGTTCATCTA 7127

QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
```

```
DB 7128 CGTCCCACCTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 7187
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 7188 GAAGCTGGACCATGGGCAACTACACACAGCTCTCTATGCCACAGAGGCCAAGTACCAC 7247
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 7248 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGTGTGCTTACCCACAGCCCTTCGCCTCC 7307
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 7308 CAGAACTGTGGCTCTGTGTGGTGGATAGTGACTATGGACCTAGACCCCAACAAGAGT 7367
QY 141 GluMetTyrAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValG1 160
DB 7368 GAATGTGGGATGTCTTCTGCTATCGGATGAAAG----- 7401
QY 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePr 180
DB 7401 ----- 7401
QY 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
DB 7402 -----GAAGTGTGGCTTATTCACACAGCTCAGCTCAGGCCG 7439
QY 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
DB 7440 TGCATTTCTAGAACACCTGACTGACCTGCTCCATCCGGCGGACCTCTTTGTGCCACAGAA 7499
QY 220 nSerGlyLeuGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
DB 7500 CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGACATCGAGCACCACTCGCCAA 7559
QY 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1 260
DB 7560 TGTGAGCATGTTTTTCTACAAAGACCTGTGTCAATGGACCACTCGCTGCAACAGAGCTGG 7619
QY 260 ySerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
DB 7620 AAGCAAGCTGCTCATCTGCTCCAGCAGGACCCACTCCACCGGAGGACCCAGTTTGT 7679
QY 280 lAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisVal11 300
DB 7680 TGATGGGAAGAGCCATCTGTCAGTGGGACATCTTTGCCCTCCAATGGGATCATTCATGTCAT 7739
QY 300 eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
DB 7740 TTCAGGCGCTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTTGGGAGC 7799
QY 320 *GlyIlePhe*****IleLeuLeuValThrGlyValAlaLeuAlaLeuAlaTyrSerTy 340
DB 7800 AGGATCTTCTTGGCCATCATCTGTCAGTGGGCTGTTCCTTGGCTGGCTTACTCTACTCTA 7859
QY 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
DB 7860 CTTTGGATAAAACCGGAGACATCGGCTTCCAGCATTTT 7999

RESULT 7
US-10-107-782-3
; Sequence 3, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Coleman, Steve
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
```



```

Db 7680 TGATGAAAGACCAATCTCGATGGGACATCTTTGCCTCCAAATGGGATCATTCATGTCAT 7739
Qy 300 eSerArg**LeuLysAlaProAlaProAlaProValThrLeu**HisThrGlyLeuGly** 320
Db 7740 TTCAGGCGCTTTAAAGACACCCCTGCCCCGTGACCTTGACCCACACTGGCTTGGGAGC 7799
Qy 320 *GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIleTyrSerTy 340
Db 7800 AGGGATCTCTTTGGCATCATCTCGTGGTGGGCTGTTCCTTGGCTGCTTACTCCTA 7859
Qy 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7860 CTTTCGGATAAACCGGAGACAAATCGGCTTCAGCATTTT 7899

RESULT 8
US-10-028-248A-1
; Sequence 1, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163)..(168)
; OTHER INFORMATION: wherein n is A, C, G, or T
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1279)..(1284)

```

```

; OTHER INFORMATION: wherein n is A, C, G, or T
US-10-028-248A-1
Alignment Scores:
Pred. NO.: 7.44e-201 Length: 8444
Score: 1554.00 Matches: 296
Percent Similarity: 85.27% Conservative: 5
Best Local Similarity: 83.85% Mismatches: 16
Query Match: 83.32% Indels: 36
DB: 15 Gaps: 1

US-09-466-778B-11 (1-353) x US-10-028-248A-1 (1-8444)
Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 6907 ATGACAGGCCCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 6966
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 6967 AACTGTGAGCGGAGCAGCTGCCCATTTGACCGCTGTTACAGGACAAATGGGACGTGCCAT 7026
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 7027 GCAGACGCCAAATGTGTGCACCTCCACTTCCAGGATACCACTGTGGGGTGTTCATCTA 7086
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 7087 CGCTCCCGACCTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCGCTGTGCCAAC 7146
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 7147 GAAGCTCGCACCATGGCAACCTCAACACAGCTCTCTATGCCCAAGAGGCCAAGTACCAC 7206
Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 7207 CTGTGCTCAGCAGCGCTGGTGGACCGGGGGGTGGCTTACCACCAAGCTTCCGCTCC 7266
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 7267 CAGAACTGTGGCTCTGGTGTGGTATGATGAGCTATGAGCTATGAGCTATGAGCTATGAGCT 7326
Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 7327 GAAATGTGGGATGTCTTGTCTATCGGATGAAG----- 7359
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 7359 ----- 7359
Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 7360 -----GAAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGCGGT 7398
Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 7399 GCATTTCTAGAACACCTGACTGACCTGTCCATCCGGGACCCCTTTGTGCTCCACAGAAC 7458
Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240
Db 7459 AGTGGCTGGGGAGAGATGAGACCTTGTCTGGGGGGGACATCGAGCACCACTCGCCCAAT 7518
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 7519 GTCAGCATGTTTTTCTACATGACCTTGTCTCAATGGCACCACTTCGCAACAGAGGCTGGGA 7578
Qy 261 SerLysLeuLeuIleThrAspArgGluAspProLeuHisProThrGluThrArgCysVal 280
Db 7579 AGCAAGCTGCTCATCTACCTGCCAGCAGGACCACTCCAAACCGGTACAAAGTAGTTGTT 7638
Qy 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 7639 GATGGAAGAGCCATCTTCGACGTGGGACATCTTTGCCCTCCAATGGGATCATTCATGTCATT 7698

```





Qy	1	MetThrGlyProGlyIysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu	20
Db	6907	ATGACAGCCCGGCGCAAGACAAAGTGTGAGTGTAAAGTCACTATGTTCGGAGATGGGCTG	6966
Qy	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
Db	6967	AACGTGAGCGGGAGCAGCTGCCCATTTGACCGCTGCTTACAGACAAATGGGCAGTGCCAT	7026
Qy	41	AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu	60
Db	7027	GCAGAGCCCAATGTGTGCGACCTCCACTTCAGGATACCACTGTGGGTGTTCATCTA	7086
Qy	61	ArgSerProLeuGlyGlnTyrLysLeuThrPheAspIysAlaArgGluAlaCysAlaAsn	80
Db	7087	CGCTCCCACTGGGCCAGTAGTAAGCTGACCTTTGACAAAGCCAGAGGGCTGTGCCAAC	7146
Qy	81	GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis	100
Db	7147	GAAGTCGGACCATGGCAACCTACAACCAAGCTCTCTATATCCCAAGGCCAAGTACCAC	7206
Qy	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaIaTyrProThrAlaPheAlaSer	120
Db	7207	CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGGTGCCTACCCCAAGCAGCTTCGGCTCC	7266
Qy	121	GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	7267	CAGAACTGTGGCTCTGTGTGTGGATAGTAGACTATGGACCTAGACCACCAACAGAGT	7326
Qy	141	GluMetTrpAspValPheCysTyrArgMetIysAspValAsnCysThr***LysValGly	160
Db	7327	GAATGTGGGATGTCCTCTGCTATCGGATGAAG-----	7359
Qy	161	TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro	180
Db	7359	-----	7359
Qy	181	SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg	200
Db	7360	-----GAAGTGCTGGCCCTATTCCAACAGCTCAGCTCAGAGGCCGT	7398
Qy	201	AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn	220
Db	7399	GCATTTCTAGAACACCTGACTGACCTGTCAATCGCGGCCACCTCTTTGTGCCACAGAAC	7458
Qy	221	SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn	240
Db	7459	AGTGGGCTGGGGAGATGAGACCTGTCTGGGGGGACATCGAGCACCACTCGCCAAT	7518
Qy	241	ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly	260
Db	7519	GTCAGCATGTTTTTCTCAATGACCTGTCAATGGCACCACTTCGAAACGAGCGTGGGA	7578
Qy	261	SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGlnThrArgCysVal	280
Db	7579	AGCAAGCTGTCTATCATCTGCGCAGCAGACCCACTCCAACGGGTACAAAGTAGGTGTGTT	7638
Qy	281	AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle	300
Db	7639	GATGGAGAGGCCATTCTGCGAGTGGGACATCTTTGCCCTCCAATGGGATCATTCATGTCATT	7698
Qy	301	SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly***	320
Db	7699	TCCAGGCGCTTTAAAGACACCCCTGCCCCCGTGAACCTTGACCCACACTGCTGTTGGGAGCA	7758
Qy	321	GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIaTyrSerTyr	340
Db	7759	GGGATCTCTTTTGCATCATCTGTGACTGGGGCTGTGGCTTGGCTGCTTACTCTCTAC	7818
Qy	341	PheArgIleAsnArgLysThrIleGlyPhe***HisPhe	353
Db	7819	TTTCGGGTAAACCGGAGACCAATCGGTACCGAGCATTTT	7857

```

US-09-842-930A-1
; Sequence 1, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 4706
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-842-930A-1

Alignment Scores:
Pred. No.: 7,57e-196 Length: 4706
Score: 1514.50 Matches: 280
Percent Similarity: 87.54% Conservative: 29
Best Local Similarity: 79.32% Mismatches: 43
Query Match: 81.21% Indels: 1
DB: 9 Gaps: 1

US-09-466-778B-11 (1-353) x US-09-842-930A-1 (1-4706)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 3022 ATGACGGGGCCAGCAGCAGCAGTAAAGTGAATGATAAAGTCACTATGTCGGGACGGAGTG 3081

Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 3082 GACTGTGAGCGCTGAGCAGCTGCGCGTTCGACCGTTCCTTACAGGACAAAGCAGCAGTGCAC 3141

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 3142 CCAGATCCAGCTGTGAGACCTCTACTTCCAGACACGACCGTAGGAGTATCCATCTA 3201

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 3202 CGCTCCCACTGGGCCAGTCAAACTGACATTTGACAAAGCAAGAGCCTGTGCCAAA 3261

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 3262 GAAGCTCGACCATGAGCCACTCAACACAGCTCTCTATGCCACAGAGGCCAGATATCAC 3321

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 3322 CTGTGCTCGCGGCTGGCTGGAGAGTGGGGGGGTTGCCTACCCGACTACGATGCTCTCT 3381

Qy 121 GlnAsnCysGlySerGlyValValGlyValLeuValAspTyrGlyProArgProAsnLysSer 140
Db 3382 CAGAAGTGTGTGCAACAGTTGTGTGGATCGTAGACTACCGATCCAGGGCCACCAAGAGT 3441

Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 3442 GAAATGTGGAGTCTCTCTGTACCGGATGAAGATGTGAAGTGCACCTTCAGAGCAGGC 3501

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 3502 TATGTGGAGATGCTCTCTCTGTGAGTGGGAACTCTGTGAGGAGTCTCTCTCTCTCTCT 3561

Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 3562 TCGGTCACAAACTTCCTGACAGAGGTGCTGGCTTTTCCAGAGCTCAGCCCGAGGACAG 3621

Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220

```



```

; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1320
; LENGTH: 3681
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1320

Alignment Scores:
Pred. No.: 4,04e-79 Length: 3681
Score: 661.50 Matches: 145
Percent Similarity: 53.44% Conservative: 49
Best Local Similarity: 39.94% Mismatches: 148
Query Match: 35.47% Indels: 21
DB: 15 Gaps: 7

US-09-466-778B-11 (1-353) x US-10-264-237-1320 (1-3681)

Qy 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
Db 2235 ACCGGCCTGAACACACACCGCGCTGTGAGTGCCACGAGGCTAGCTAGGCGATGGACTGCAG 2294

Qy 22 Cys---GluProGluGlnLeuProLeuAspArgCysGluGlnAspAsnGlyGlnCysHis 40
Db 2295 TGTCTGGAGGAGTGGACACCGCTGTGACCGCTCTTGGGCGCAGCCACCGCCCTGCGCAC 2354

Qy 41 AlaPheAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 2355 TCAGATGCCATGTGCACTGACCTGACCTGCCAGGAGAAACGGGCTGGCGTTTTCACCTC 2414

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaAargGluAlaCysAlaAsn 80
Db 2415 CAGGCCACACGCGGCCCTTATGCTGAACCTTCGGAGGCTGAGCGGCATGGGAACA 2474

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 2475 CAGGAGCGCGCTGTGCTTCATCCCTCAGCTCTCTGCTGCCACGAGCTGGGCTTCAC 2534

Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 2535 CTGTGCTCATGGGCTGTGGCTGCGCAATGGCTCCACTGCCACCGCTGTGCTTCCCTGTG 2594

Qy 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
Db 2595 GCGGATGTGGCAATGGTGGGTCATGTCAGCTGTCAGCTGGGCGCCGCAAGAACCTCTCA 2654

Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 2655 GAACGCTGGATGCTACTGCTCTGCTGTGCGAAGATGGCGCTGCCGATGCCGAATGGC 2714

Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnValLeuMetSerPhe 179
Db 2715 TTCGTGGGTGACGGGATGACGACCGTGAATGGGAAGCTGCTGGATGTGCTGCTGCCACT 2774

```

```

Qy 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199
Db 2775 GCCAACTTCTCCACCTTCTATGGGATGCTATGGGCTATGCCATGCCACCGCGGGT 2834

Qy 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerLeuArgGlyThrLeuPheValProGln 219
Db 2835 CTGACTTCTGGACTTCTGGATGCTGAGTACAGTATAGACACTCTTCGTCCTCCCTGC 2894

Qy 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
Db 2895 AATGAAGGCTTTGFGGACAAACATGAGCTGAGTGGCCAGACTTGGAGCTGCATGCCCTC 2954

Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
Db 2955 AACGCCACCTCTTAAGTGCCAAAC---GCCAGCCAGGGGAAGTTGCTTCCGGCCCACTCA 3011

Qy 260 GlySerLysLeuLeuIleThrAspArg-----GlnAspProLeuHisProThrGlu 276
Db 3012 GGCCTCAGCTCATCATCAGTACGAGGAGCCCTCACACAGTTCCTGGGCCCTGTGGCC 3071

Qy 277 ThrArgCysValAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIle 296
Db 3072 CCAGGACAGTTGTGTTAGCGGTATCATTTGTGGGACATCATGGCTTCAATGGCATC 3131

Qy 297 ThrHisValIleSerArg***LeuLysAlaProProAlaProValThrLeu----- 313
Db 3132 ATCCATGCTCTGGCCAGCCCTCTCTGGCACCCCCACAGCCCGGAGGCTGTGGCGCT 3191

Qy 314 -----***HisThrGlyLeuGly***GlyIlePhe***IleIleLeuVal 329
Db 3192 GAAGCCCCACCTGTGGCGGAGGCTGTGGGCT-----GTGCTTGGC 3233

Qy 330 ThrGlyAlaVal-----AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgLysThr 347
Db 3234 GCTGAGCACTGTGCTGTGGTGGCGGAGCTCTACTCTCCCTGCCGCGAGGCAAGCC 3293

Qy 348 IleGlyPhe 350
Db 3294 ATGGCGTTT 3302

RESULT 13
US-09-774-639-12
; Sequence 12, Application US/09774639
; Publication No. US20030003555A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-774-639-12

Alignment Scores:
Pred. No.: 3.96e-77 Length: 2753
Score: 645.50 Matches: 143
Percent Similarity: 52.89% Conservative: 49
Best Local Similarity: 39.39% Mismatches: 150
Query Match: 34.61% Indels: 21
DB: 10 Gaps: 7

US-09-466-778B-11 (1-353) x US-09-774-639-12 (1-2753)

Qy 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
Db 264 ACCGGCCTGAACACACCGCGCTGTGAGTGGCGGCTAGCTAGGCTAGGCTAGGCTAGG 323

```

Qy	22	Cys---GluProGluInLeuProileAspArgCysLeuGlnAepAenglyGlnCysHis	40
Db	324	TGTCGTGGAGAGTCCGAACACACCTGTGTGAACCGCTGCTTGGCGCAGCCACCGCCCTGCCAC	383
Qy	41	AlaAspAlaIysCysValAspLeuHiPheGlnAspThrThrValGlyValPheHisLeu	60
Db	384	TCAGATGCCATGTGMACTGCACCTGCATCTCCAGAGAAACGGGTGGCGTGTTCACCTC	443
Qy	61	ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	444	CAGGCCACCGCGCCCTTATGTGCTGAACCTTTTCGGAGGCTGAGCGCGCATCGAAGCA	503
Qy	81	GluAlaIaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaIaYsYrHis	100
Db	504	CAGGAGCGCGCTTGTGCTTCACTCCCTCAGCTCTCTGTCGCCACGACGTGGGTCTCCAC	563
Qy	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaIaTyrProThrAlaPheAlaSer	120
Db	564	CTGTGCTCATGGCTGGCTGGCAATGGCTCCACTGCCACCCTGTGGTTTCCCTGTG	623
Qy	121	GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	624	CGCGACTGTGGCAATGTGTGGGTGGGCTAGTCAGCCTGGGTGGCGCCCAAGAACCTCTCA	683
Qy	141	GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly	160
Db	684	GAACGCTGGGATGGCTACTGTCTCCGTGTGCAAGATGTGGCTGCCGATGCCGAATGGC	743
Qy	161	TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuGlnValLeuMetSerPhe	179
Db	744	TTCTGGGTGACGGATCAGACAGCTGCATGGGAAGCTGTGGATGTCTGGTGGCTCACT	803
Qy	180	ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly	199
Db	804	GCCAACTTCTCCACTTCTATGGGATGCTATTGGGCTATGCCAATGCCACCCAGCGGGT	863
Qy	200	ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln	219
Db	864	CTCGACTTCCTGGACTTCTCGTGGATGATGAGCTACGTATAAGACACTCTTCGTCCCTGTC	923
Qy	220	AsnSerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisHisLeuAla	239
Db	924	AATGAAGCTTTGTGGACAACATGACCTGAGTGGCCCCAGACTTGGAGCTGCATGCCCTCC	983
Qy	240	AsnValSerMetPhePheTyrAsnAspLeuValAenglyThrThrLeuGlnThrArgLeu	259
Db	984	AACGCCACCTCTCAAGTGCCAAC---GCCAGCCAGGGGAAGTGTCTTCGGCCCACTCA	1040
Qy	260	GlySerLysLeuLeuIleThrAspArg-----GlnaspProLeuHisProThrGlu	276
Db	1041	GGCTCAGCCTCATCATCAGTGCAGCGCGCCCTGCACAACTTCCCTGGGCCCCCTGTGGCC	1100
Qy	277	ThrArgCysValAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIle	296
Db	1101	CCAGGGACATGTGGTTAGCCGTATCATTTGTGGACATCATGGCTTCAATGGCATC	1160
Qy	297	ThrHisValIleSerArg***LeuLysAlaProProAlaProValThrLeu-----	313
Db	1161	ATCCATGTCTTGGCGAGCCCCCTCTGTGCACCCACAGGCCACAGGCAGTGTGGCGCT	1220
Qy	314	-----***HisThrGlyLeuGly***GlyIlePhe****IleIleLeuVal	329
Db	1221	GAAGCCCCACTGTGTGGCGGAGGGCTGTGGGGCT-----GTGCTTGGCC	1262
Qy	330	ThrGlyAlaVal-----AlaLeuAlaIaTyrSerTyrPheArgIleAsnArgLysThr	347
Db	1263	GCTGGAGCACTGTGGTGTGGTGGCGGAGCTCTCTACCTCCGTCCGTGCCGAGGCAAGCCC	1322
Qy	348	IleGlyPhe	350
Db	1323	ATGGCTTT	1331

```

RESULT 14
US-09-969-730-13
; Sequence 13, Application US/09969730
; Publication No. US2003005443A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: PZ013P2
; CURRENT APPLICATION NUMBER: US/09/969,730
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/056,371
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,367
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,365
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,731
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,557
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,563
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/055,970
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,986
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,311
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,808
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,803
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,809
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,806
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,310
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,798
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,309
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,312
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,807
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,386
; PRIOR FILING DATE: 1997-08-05
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-730-13

```

Alignment Scores:  
Pred. No.: 3,96e-77 Length: 2753  
Score: 645.50 Matches: 143  
Percent Similarity: 52.89% Conservative: 49  
Best Local Similarity: 39.39% Mismatches: 150  
Query Match: 34.61% Indels: 21  
DB: 10 Gaps: 7

US-09-466-778B-11 (1-353) x US-09-969-730-13 (1-2753)

Qy 2 ThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21  
Db 264 ACCGCCCTGAACACACACCGCGCTGTAGTGCACCGAGCTACGTAGGCGATGGAGCTGCAG 323

Qy 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
Db 324 TGTCTGGAGGATCGAACCACCTGTGACCGCTCTGGGCCAGCCAGCCCGCTGCCAC 383

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
Db 384 TCAGATGCATGTGMACTGCATCTCCAGGAGAAACGGCGCTGGCGTTTTCACCTC 443

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
Db 444 CAGGCCACAGCGCCCTTATGCTGTAACCTTCGAGGCTGAGCGGCATCGCAAGCA 503

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
Db 504 CAGGAGCGCTCTGCTTCATTCCTCAGCTCTCTGCTGCCACGAGCTGGCTTCCAC 563

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
Db 564 CTGTGCTCATGGCTGGCTGGCCAAATGGCTCCACTGCCACCTGTGTTTTCCTCTGT 623

Qy 121 GlnAsnGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140  
Db 624 GCGGACTGGCAATGGCTGGGTGGCTAGTACCTGGGTGGCGCCGCAAGAACTCTCA 683

Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
Db 684 GAAAGCTGGATGCTACTGCTGCTGCTGCAAGATGGCTGCGCATGCCGAATGGC 743

Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnValLeuMetSerPhe 179  
Db 744 TTCGTGGGTGACGGATCAGCACGTGCAATGGAAAGCTGCTGATGTGCTGCTGCCACT 803

Qy 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerAlaArgGly 199  
Db 804 GCCAACTCTCCACCTTCTATGGGATGCTATTTGGGCTATGCCAATGCCACCGGGGT 863

Qy 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219  
Db 864 CTCGACTTCTGGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 923

Qy 220 AsnSerGlyLeuGlyLysAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239  
Db 924 AATGAAGCTTTGTGGACAACTATGCTGATGATGATGATGATGATGATGATGATGATGAT 983

Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259  
Db 984 AACGCCACCTCTTAAGTGCAC---GCCACCGAGGGGAAGTTGCTCCGCCCACTCA 1040

Qy 260 GlySerLysLeuLeuLeuThrAspArg-----GlnAspProLeuHisProThrGlu 276  
Db 1041 GGCCTCAGCCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100

Qy 277 ThrArgCysValAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIle 296  
Db 1101 CCAGGACAGTTGTGGTTAGCCGTATCATTTGTGGGACATCATGCGCTTCAATGGCATC 1160

Qy 297 ThrHisValIleSerArg\*\*\*LeuLysAlaProAlaProValThrLeu----- 313

Db 1161 ATCATGCTCTGGCCAGCCCTCTCTGGACACCCCAAGCCAGGAGCTGCTGGCGCT 1220

Qy 314 -----\*\*\*HisThrGlyLeuGly\*\*\*GlyIlePhe\*\*\*\*\*IleIleLeuVal 329  
Db 1221 GAAGCCCACTGTGGCGGAGCGGTGGGGCT-----GTGCTTGCC 1262

Qy 330 ThrGlyAlaVal-----AlaLeuAlaTyrSerTyrPheArgIleAsnArgLysThr 347  
Db 1263 GCTGAGCACTGTGCTGGCTGGTGGCGGAGCTCTACCTCGTGGCCGCAAGGCC 1322

Qy 348 IleGlyPhe 350  
Db 1323 ATGGCTTT 1331

RESULT 15  
US-10-621-363-13  
; Sequence 13, Application US/10621363  
; Publication No. US20040023283A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P2C1  
; CURRENT APPLICATION NUMBER: US/10/621,363  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 09/969,730  
; PRIOR FILING DATE: 2001-10-06  
; PRIOR APPLICATION NUMBER: 09/774,639  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/238,291  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: PCT/US98/16235  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/056,371  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,732  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,366  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,364  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 373  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 2753  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-621-363-13

Alignment Scores:  
Pred. No.: 3,96e-77 Length: 2753  
Score: 645.50 Matches: 143  
Percent Similarity: 52.89% Conservative: 49  
Best Local Similarity: 39.39% Mismatches: 150  
Query Match: 34.61% Indels: 21  
DB: 10 Gaps: 7

US-09-466-778B-11 (1-353) x US-10-621-363-13 (1-2753)

Qy 2 ThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21  
Db 264 ACCGCCCTGAACACACACCGCGCTGTAGTGCACCGAGCTACGTAGGCGATGGAGCTGCAG 323

Qy 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
Db 324 TGTCTGGAGGATCGAACCACCTGTGACCGCTCTGGGCCAGCCAGCCCGCTGCCAC 383

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
Db 384 TCAGATGCATGTGMACTGCATCTCCAGGAGAAACGGCGCTGGCGTTTTCACCTC 443

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
Db 444 CAGGCCACAGCGCCCTTATGCTGTAACCTTCGAGGCTGAGCGGCATCGCAAGCA 503

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
Db 504 CAGGAGCGCTCTGCTTCATTCCTCAGCTCTCTGCTGCCACGAGCTGGCTTCCAC 563

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
Db 564 CTGTGCTCATGGCTGGCTGGCCAAATGGCTCCACTGCCACCTGTGTTTTCCTCTGT 623

Qy 121 GlnAsnGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140  
Db 624 GCGGACTGGCAATGGCTGGGTGGCTAGTACCTGGGTGGCGCCGCAAGAACTCTCA 683

Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
Db 684 GAAAGCTGGATGCTACTGCTGCTGCTGCAAGATGGCTGCGCATGCCGAATGGC 743

Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnValLeuMetSerPhe 179  
Db 744 TTCGTGGGTGACGGATCAGCACGTGCAATGGAAAGCTGCTGATGTGCTGCTGCCACT 803

Qy 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerAlaArgGly 199  
Db 804 GCCAACTCTCCACCTTCTATGGGATGCTATTTGGGCTATGCCAATGCCACCGGGGT 863

Qy 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219  
Db 864 CTCGACTTCTGGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 923

Qy 220 AsnSerGlyLeuGlyLysAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239  
Db 924 AATGAAGCTTTGTGGACAACTATGCTGATGATGATGATGATGATGATGATGATGATGAT 983

Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259  
Db 984 AACGCCACCTCTTAAGTGCAC---GCCACCGAGGGGAAGTTGCTCCGCCCACTCA 1040

Qy 260 GlySerLysLeuLeuLeuThrAspArg-----GlnAspProLeuHisProThrGlu 276  
Db 1041 GGCCTCAGCCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100

Qy 277 ThrArgCysValAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIle 296  
Db 1101 CCAGGACAGTTGTGGTTAGCCGTATCATTTGTGGGACATCATGCGCTTCAATGGCATC 1160

Qy 297 ThrHisValIleSerArg\*\*\*LeuLysAlaProAlaProValThrLeu----- 313

Db 384 TCAGATGCCATGTGMACTGACCTGCACCTCCAGGAGAAACGGGCTGGCGTTTCCACCTC 443  
QY : : : : :  
Db 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
: : : : :  
Db 444 CAGGCCACACAGCGCCCTATATGCTCTGAACCTTTTCGGAGCTGAGGCGCATCGAAGCA 503  
: : : : :  
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
: : : : :  
Db 504 CAGGAGCGCTCCTCTGCTTCCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCAC 563  
: : : : :  
QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
: : : : :  
Db 564 CTGTGCTCATGGGCTGGCTGGCCATGGCTCCACTGCCCCCTGTGGTTTCCCTGTG 623  
: : : : :  
QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
: : : : :  
Db 624 GCGGACTGTGGCAATGCTCGGGTGGGCTAGTCAGCGCTGGGTGCCCGCAAGAACCTCTCA 683  
: : : : :  
QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
: : : : :  
Db 684 GAACGCTGGGATGCCTACTCTCCGTGTGCAAGATGTGGCTGCCGATGCCGAAATGGC 743  
: : : : :  
QY 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnValLeuMetSerPhe 179  
: : : : :  
Db 744 TTCGTGGTGACGGGATCAGCACGTGCATGGGAAGCTGTGGATGTGCTGGGTGCCACT 803  
: : : : :  
QY 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199  
: : : : :  
Db 804 GCCAATCTCTCCACTTCTATGGGATGCTATTTGGGCTATGCCAATGCCACCCAGCGGGT 863  
: : : : :  
QY 200 ArgAlaPheLeuGlnHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219  
: : : : :  
Db 864 CTCGACTTCTGAGCTTCTCGGATGTAGTACGTATAGACACTCTTCGTCCCTGTC 923  
: : : : :  
QY 220 AsnSerGlyLeuGlyGlnAsnGluThrLeuSerGlyArgAspIleGlnHisLeuAla 239  
: : : : :  
Db 924 AATGAGGCTTTTGGACACATCAGCTGAGTGGCCACAGCTTGGAGCTGCATGCCCTCC 983  
: : : : :  
QY 240 AsnValSerMetPheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259  
: : : : :  
Db 984 AACGCCACCTCTTAAGTGCAC---GCCAGCCAGGGGAAGTTGCTTCGGGCCACTCA 1040  
: : : : :  
QY 260 GlySerLysLeuLeuIleThrAspArg-----GlnAspProLeuHisProThrGlu 276  
: : : : :  
Db 1041 GGCCTCAGCCTCATCATCAGTACGACGAGCGCCCTGACACAGTTCTCTGGGCCCTGTGGCC 1100  
: : : : :  
QY 277 ThrArgCysValAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIle 296  
: : : : :  
Db 1101 CCAGGGACAGTTGGTTAGCGGTATCATTTGTGGGACATCATGGCTTCATATGGCATC 1160  
: : : : :  
QY 297 ThrHisValIleSerArg\*\*\*LeuLysAlaProProAlaProValThrLeu----- 313  
: : : : :  
Db 1161 ATCCATGCTCTGGCCAGCCCTCTCTGGCACCCCTCCAGCCACAGCCAGGCGAGTGTGGCGCCT 1220  
: : : : :  
QY 314 -----\*\*\*HisThrGlyLeuGly\*\*\*GlyIlePhe\*\*\*\*\*IleLeuVal 329  
: : : : :  
Db 1221 GAGCCCCACTGTGGCGGAGGGGTGGGGCT-----GTGCTTGCC 1262  
: : : : :  
QY 330 ThrGlyAlaVal-----AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgLysThr 347  
: : : : :  
Db 1263 GCTGGAGCACTGTGCTTGGCTTGGTGGCGGAGCTCTTACCTCGTCCGCGGCAAGCCC 1322  
: : : : :  
QY 348 IleGlyPhe 350  
: : : : :  
Db 1323 ATGGGCTTT 1331  
: : : : :

Search completed: April 1, 2004, 14:03:06  
Job time : 1541 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 121 Seconds  
(without alignments)  
1618.989 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865  
Sequence: 1 MTGPGHKCKSHYVDGL.....ALAAYSYFRINRKTIGXHF 353

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Xgapop 10.0, Xgapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09466778/runat\_01042004\_095643\_10608/app\_query.fasta\_1.519  
-DB=Issued\_Patents\_NA -QPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DILOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09466778 @CIGN 1 1 69 @runat\_01042004\_095643\_10608 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632.5	33.9	1482	4	US-09-016-434-230
2	266.5	14.3	1414	1	US-08-024-868-1
3	266.5	14.3	1414	2	US-08-242-097-1
4	266.5	14.3	1414	3	US-09-206-695-1
5	266.5	14.3	1414	4	US-09-799-118-1
6	255.5	13.7	1734	4	US-09-484-970B-63
7	200.5	10.8	1985	4	US-09-907-794A-212
8	200.5	10.8	1985	4	US-09-905-125A-212
9	200.5	10.6	1985	4	US-09-902-775A-212
10	198.5	10.6	409	4	US-09-833-381-1252
11	157.5	8.4	1520	1	US-08-225-477B-1
12	157.5	8.4	1520	5	PCT-US95-04353-1

13	157.5	8.4	5191	1	US-08-340-428B-1	Sequence 1, Appli
14	157.5	8.4	5191	5	PCT-US93-07306-1	Sequence 1, Appli
15	155.5	8.3	3259	5	PCT-US95-03747-1	Sequence 1, Appli
16	155	8.3	1587	4	US-09-010-147B-19	Sequence 19, Appli
17	154.5	8.3	1400	2	US-08-001-078A-2	Sequence 2, Appli
18	154.5	8.3	1400	2	US-08-463-218-2	Sequence 2, Appli
19	154.5	8.3	1400	5	PCT-US94-00253-2	Sequence 2, Appli
20	153.5	8.2	1519	1	US-08-225-477B-2	Sequence 2, Appli
21	153.5	8.2	1519	5	PCT-US95-04353-2	Sequence 2, Appli
22	152.5	8.2	1720	4	US-09-148-545-53	Sequence 53, Appli
23	150.5	8.1	8224	6	5180808-1	Patent No. 5180808
24	122.5	6.6	1962	4	US-09-148-545-111	Sequence 111, App
25	122	6.5	1896	4	US-09-724-864-28	Sequence 28, Appli
26	109	5.8	2029	4	US-09-232-160-13	Sequence 13, Appli
27	109	5.8	2372	4	US-09-907-794A-200	Sequence 200, App
28	109	5.8	2372	4	US-09-905-125A-200	Sequence 200, App
29	109	5.8	2372	4	US-09-902-775A-200	Sequence 200, App
30	109	5.8	2404	4	US-09-833-381-849	Sequence 849, App
31	109	5.8	2404	4	US-09-833-381-853	Sequence 853, App
32	107	5.7	2313	2	US-08-892-880-1	Sequence 1, Appli
33	105	5.6	2430	2	US-08-820-170A-35	Sequence 35, Appli
34	105	5.6	2430	3	US-09-055-699-35	Sequence 35, Appli
35	105	5.6	2430	3	US-09-273-565-35	Sequence 35, Appli
36	105	5.6	2430	4	US-09-565-538-35	Sequence 35, Appli
37	105	5.6	2430	4	US-09-661-468-35	Sequence 35, Appli
38	105	5.6	2430	4	US-09-976-165-35	Sequence 35, Appli
39	105	5.6	2977	2	US-08-820-170A-36	Sequence 36, Appli
40	105	5.6	2977	3	US-09-055-699-36	Sequence 36, Appli
41	105	5.6	2977	3	US-09-273-565-36	Sequence 36, Appli
42	105	5.6	2977	4	US-09-565-538-36	Sequence 36, Appli
43	105	5.6	2977	4	US-09-661-468-36	Sequence 36, Appli
44	105	5.6	2977	4	US-09-976-165-36	Sequence 36, Appli
45	101.5	5.4	3207	1	US-07-946-497-1	Sequence 1, Appli

## ALIGNMENTS

### RESULT 1

US-09-016-434-230  
; Sequence 230, Application US/09016434  
; Patent No. 6500938

; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555



```

RESULT 2
US-08-024-868-1
: Sequence 1, Application US/08024868
: Patent No. 5386013
: GENERAL INFORMATION:
: APPLICANT: Lee, Tae Ho
: APPLICANT: Wisniewski, Hans Georg
: APPLICANT: Vilcek, Jan
: TITLE OF INVENTION: Cytokine-Induced protein, TSG-6, DNA
: TITLE OF INVENTION: Coding Therefor and Uses Thereof
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Browdy and Neimark
: STREET: 419 Seventh Street, NW
: CITY: Washington
: STATE: DC
: COUNTRY: US
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/024,868
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/642,312
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Livnat, Shmuel
: REGISTRATION NUMBER: 33,949
: REFERENCE/DOCKET NUMBER: VILCEK-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-628-5197
: TELEFAX: 212-737-3528
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1414 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: CELL TYPE: Fibroblast
: CELL LINE: FS-4
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 69..899
: OTHER INFORMATION:

```





```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1414 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   ORIGINAL SOURCE:
;     ORGANISM: Homo sapiens
;     CELL TYPE: Fibroblast
;     CELL LINE: FS-4
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 69..899
;   OTHER INFORMATION:
;     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-799-118-1

```

```

Alignment Scores:
Pred. No.: 9,99e-25 Length: 1414
Score: 266.50 Matches: 64
Percent Similarity: 47.42% Conservative: 28
Best Local Similarity: 32.99% Mismatches: 78
Query Match: 14.29% Indels: 24
DB: 4 Gaps: 4

```

US-09-466-778B-11 (1-353) x US-09-799-118-1 (1-1414)

```

Qy 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPhe 71
Db 162 GAACGAGCAGCGGTGTGTACACAGAGAGCAGCGTCTGGCAAAATACAACTCACCTAC 221
Qy 72 AspLysAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGlnLeu 91
Db 222 GCAGAGCTAAGCGCGGTGTGCAATTGAAAGCGCGCCATCTCGCAACTTACAAGCAGCTA 281
Qy 92 SerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
Db 282 GAGCGACCCAGAAAATTTGGATTTCATGCTGTCTGCTGGATGGATGGCTAAGGGCAGA 341
Qy 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyLeVal 131
Db 342 GTTGGATACCCCATTTGTAAGCCAGGCGCCCACTGTGGATTGGAAAACTGGCATTATT 401
Qy 132 AspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArgMetLys 151
Db 402 GATTATGGAATCCCTCTCAATAGAGTGAAGATGGGATGGCTATTGTGTACAACCCACAC 461
Qy 152 AspValAsnCys-----Thr***LysValGlyTyrValGlyAspGlyPhe 166
Db 462 GCAAGAGGTGTGTGGCGCTTTTACAGATCCAAAGCGAAATTTTAAATCTCCAGGCTTC 521
Qy 167 -----SerTyrSerGlyAsnLeu----- 173
Db 522 CCAATAGTACGAAGATAACAAATCTGCTACTGGCACATTAGACTCAAGTATGGTCAG 581
Qy 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189
Db 582 CGTATTACCTGAGATTTTGTAGATTTTACCTTGAAGATGACCAGGTGTGCTGGCTGAT 641
Qy 190 ----LeuAlaTyrSerAsnSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAs 208
Db 642 TATGTTGAATATATGACATTGACATGATGATGCTCCATGGCTTTGGGAAGATATCTGTGGA 701
Qy 208 pLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
Db 702 GATGAGCTTCCAGATGACATCATCAGTACAGGAAATGTCA 741

```

RESULT 6

```

US-09-484-970B-63
; Sequence 63, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.

```

```

; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 63
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 201752.1
; NAME/KEY: unsure
; LOCATION: 1444, 1455, 1457, 1461, 1463, 1465, 1467, 1497, 1500, 1506, 1521, 1534-
; OTHER INFORMATION: a, c, g, or other
US-09-484-970B-63

```

```

Alignment Scores:
Pred. No.: 3.92e-23 Length: 1734
Score: 255.50 Matches: 64
Percent Similarity: 47.18% Conservative: 28
Best Local Similarity: 32.82% Mismatches: 78
Query Match: 13.70% Indels: 25
DB: 4 Gaps: 4

```

US-09-466-778B-11 (1-353) x US-09-484-970B-63 (1-1734)

```

Qy 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPhe 71
Db 162 GAACGAGCAGCGGTGTGTACACAGAGAGCAGCGTCTGGCAAAATACAACTCACCTAC 221
Qy 72 AspLysAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGlnLeu 91
Db 222 GCAGAGCTAAGCGCGGTGTGCAATTGAAAGCGCGCCATCTCGCAACTTACAAGCAGCTA 281
Qy 92 SerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
Db 282 GAGCGACCCAGAAAATTTGGATTTCATGCTGTCTGCTGGATGGATGGCTAAGGGCAGA 341
Qy 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyLeVal 131
Db 342 GTTGGATACCCCATTTGTAAGCCAGGCGCCCACTGTGGATTGGAAAACTGGCATTAT 401
Qy 131 LAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArgMetLys 151
Db 402 TGATTATGGAATCCCTCTCAATAGAGTGAAGATGGGATGGCTATTGTGTACAACCCACA 461
Qy 151 sAspValAsnCys-----Thr***LysValGlyTyrValGlyAspGlyPhe 166
Db 462 CGCAAGAGGTGTGTGGCGCTTTTACAGATCCAAAGCAAATTTTAAATCTCCAGGCTTC 521
Qy 166 e-----SerTyrSerGlyAsnLeu----- 173
Db 522 CCAATAGTACGAAGATAACAAATCTGCTACTGGCACATTAGACTCAAGTATGGTCAG 581
Qy 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189
Db 582 CGTATTACCTGAGATTTTGTAGATTTTACCTTGAAGATGACCAGGTGTGCTGGCTGTA 641
Qy 190 ----LeuAlaTyrSerAsnSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAs 208
Db 642 TTATGTTGAATATATGACAGATTACGATGATGCTCCATGGCTTTGTGGAGAGATCTGTGG 701
Qy 208 sLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
Db 702 AGATGAGCTTCCAGATGACATCATCAGTACAGGAAATGTCA 742

```

RESULT 7

```

US-09-307-794A-212
; Sequence 212, Application US/09307794A

```

```

Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavir, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 212
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-794A-212

```

Alignment Scores:

Pred. No.:	8.42e-16	Length:	1985
Score:	200.50	Matches:	84
Percent Similarity:	35.71%	Conservative:	46
Best Local Similarity:	23.08%	Mismatches:	93
Query Match:	10.75%	Indels:	141
DB:	4	Gaps:	14

US-09-466-778B-11 (1-353) x US-09-907-794A-212 (1-1985)

```

QY 5 GlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsnCysGluPro 24
Db 524 GGGCGTTACCGCTGTGAG-----GTCATTGACGGCGTCGAGGATGAA--- 565
QY 25 GluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLys 44
Db 566 -----AGCGGT 571
QY 45 CysValAspLeuHisPheGlnAspThrValGlyValPheHisLeuArgSerProLeu 64
Db 572 CTGGTGGAGCTGGAGCTGCGGGGTG-----GTCTTTCTCTTACCGATCCGCCAAC 622
QY 65 GlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaLathr 84
Db 623 GGGCGCTACCACTTCCAGAGCGCCAGGAGTCTGTGACAGCAGCTGCGGTG 682
QY 85 MetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAla 104
Db 683 GTGGCTCTCTTTGAGCAGCTCTCCGGGCTGGGAGGAGGCGCTGGACTGGTGAACGCG 742
QY 105 GlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGly 124
Db 743 GGCTGGCTGCAGATGCTACCGTACAGTACCCATCATGTTGCCCGCCGACCCCTGCGGT 802
QY 125 Ser---GlyValVal---GlyLeuValAspTyrGlyProArgProAsnLysSerGluMet 142
Db 803 GGCCAGCGCTGGCACCTGGCGTGGGAAGTACGGCCCCCGCCACCGCCGCTGCACCGC 862
QY 143 TrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGlyTyrVal 162
Db 863 TATGATGTATTCTGCTCGCT----- 883
QY 163 GlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSerLeu 182
Db 883 ----- 883
QY 183 ThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla--- 201
Db 884 -----ACTGCCCTCAAGGGCGGGTGTAC 907
QY 202 PheLeuGluHisLeuThrAspLeuSerIleArg----- 212
Db 908 TACCTGGAGCACCTCGAGAGCTGACGCTGACAGAGGCAAGGAGGCGCTGCAGGAAGAT 967
QY 213 -----GlyThrLeuPheVal-ProGlnAsnSerGlyLeuGlyG1 225
Db 968 GATGCCAGCATCGCAAGTGGGACAGCTCTTTCGGCTCGAAGTTCATGCGCTGGAC 1027
QY 225 uAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePh 245
Db 1028 CGCTCGAGCGTGGCTGGCTGGCAGATG-----GCAGCGTCCGCTACCTGTGGTT 1078
QY 245 eTyrAsnAspLeuValAsnGlyThrLeuGlnThrArgLeuGlySer-LysLeuLeu1 265
Db 1079 CACCGCATCTCTAACTGTGGGC-----CCCCAGAGCTGGGTCCGAAGCTTGGC 1129
QY 265 leThrAspArgGlnAspProLeuHisProThrGluThrArgCysValAspGlyArgSept 285
Db 1130 TTCCCCGACCCGACAGCCGCTT-----GTACGGTGT----- 1161
QY 285 hrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg***LeuL 305
Db 1162 -----TTAC 1165

```



```
QY 225 uAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePh 245
Db 1028 CGTGCAGCGTGGCTGGCTGCAGATG-----GCAGGCTCCGCTACCTGTGGTT 1078
QY 245 eTyraAsnAspLeuValAlaGlyThrLeuGlnThrArgLeuGlySer-LysLeuLeu 265
Db 1079 CACCCGATCTTACTGTGGG-----CCCCAGAGCTGGGGTCCGAGCTTTGGC 1129
QY 265 leThrAspArgGlnAspProLeuHisProThrGluThrArgCysValAspGlyArgAsp 285
Db 1130 TTCCCGAGCCCGCAGAGCGCTT-----GTACGGTGT-----1161
QY 285 hrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg***Leu 305
Db 1162 -----TTAC 1165
QY 305 ysAlaProAlaProValThrLeu***HisThrGlyLeuGly***GlyIlePhe----- 323
Db 1166 TGCTACCCGACCA-----CTAGGACCTGGGCCCCCTCCCT 1201
QY 324 -----*****lelleLeuValThrGlyAlaValAlaLeuAlaIleYrs 339
Db 1202 GCGCATTCCTCACTGCTGTGTATTATTAGAGGTTCGTTTCCCTTGTGGTTGGA 1261
QY 339 erTyPhe 341
Db 1262 GCCATTIT 1269

RESULT 9
US-09-902-775A-212
; Sequence 212, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Giang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary B.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
```

```
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-775A-212

Alignment Scores:
Pred. No.: 8,42e-16 Length: 1985
Score: 200.50 Matches: 84
Percent Similarity: 35.71% Conservative: 46
Best Local Similarity: 23.08% Mismatches: 93
Query Match: 10.75% Indels: 141
DB: 4 Gaps: 14

US-09-466-778B-11 (1-353) x US-09-902-775A-212 (1-1985)
QY 5 GlyLysHisLysCysGluCysLysSerHisTyValGlyAspGlyLeuAsnCysGluPro 24
Db 524 GGGCTTACCGCTGTGAG-----GTCTTACCGGCTGGAGGATGAA--- 565
QY 25 GluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLys 44
Db 566 -----AGCGGT 571
QY 45 CysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeu 64
Db 572 CTGCTGGAGCTGGAGCTCCGGGGTGTG-----GTCTTCTTACCGTCCCAAC 622
QY 65 GlyGlnTyLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThr 84
Db 623 GGGCGCTACCACTTCCAGAGGCGCAGAGCTCTGTGAGAGAGGCTGCGGTG 682
QY 85 MetaLThrTyraAsnGlnLeuSerTyra***GlnLysAlaLysTyHisLeuCysSerAla 104
Db 683 GTGGCTTCCCTTTGAGCAGCTTCCGGGCGCTGGAGGAGGCGCTGGAGCTGGTCAACGCG 742
QY 105 GlyTrpLeuGluThrGlyArgValAlaTyProThrAlaPheAlaSerGlnAsnCysGly 124
Db 743 GGCTGGCTGAGGATGCTACGCTGAGTACCCATCATGTTGCCCGGCGAGCCCTGCGGT 802
QY 125 Ser---GlyValVal---GlyIleValAspTyGlyProArgProAsnLysSerGluMet 142
Db 803 GGCCAGGCTGGCACCTGGCGTGGAGAGTACGGCCCCCGCCACCCGCCCTGTCACCGC 862
QY 143 TrpAspValPheCysTyraArgMetLysAspValAsnCysThr***LysValGlyTyVal 162
Db 863 TATGATGATTTCTGCTTCGCT-----883
```

QY 163 GlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSerLeu 182  
 Db 883 ----- 883  
 QY 183 ThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla--- 201  
 Db 884 -----ACTGCCCTCAAGGGCGGGTGATC 907  
 QY 202 PheLeuGluHisLeuThrAspLeuSerIleArg----- 212  
 Db 908 TACCTGGAGCACCCTGAGAGCTGACCTGACAGAGGCAAGAGGAGGCTGCGAGAGAT 967  
 QY 213 -----GlyThrLeuPheVal-ProGlnAsnSerGlyLeuGlyG1 225  
 Db 968 GATGCCACGATCGCAAGGTGGGACAGCTCTTTGCCCGCTGGAGTTCCATGCCCTGGAC 1027  
 QY 225 uAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPheP 245  
 Db 1028 CGCTGCCAGCTGGCTGGCGAGATG-----GCAGCGTCCGCTACCCCTGGTT 1078  
 QY 245 eTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGlySer-LysLeuLeu 265  
 Db 1079 CACCGCATCTTAAGTGGGC-----CCCCAGCGCTGGGTCCGAGCTTTGGC 1129  
 QY 265 leThrAspArgGlnAspProLeuHisProThrGluThrArgCysValAspGlyArgAsp 285  
 Db 1130 ITCCCGACCGCGAGCCGCTT-----GTACGGTGT----- 1161  
 QY 285 hrLeuGluTfAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg\*\*\*Leu 305  
 Db 1162 -----TTAC 1165  
 QY 305 ysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\*GlyIlePhe--- 323  
 Db 1166 TGCTACCGCAGCA-----CTAGGACCTGGGGCCCTCCCT 1201  
 QY 324 -----\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrS 339  
 Db 1202 GCGCATTCCTCACTGGTGTATTATTAGTGTGCTTCCCTTTGCTGGTTTGA 1261  
 QY 339 erTyrPhe 341  
 Db 1262 GCCATTIT 1269  
 RESULT 10  
 US-09-833-381-1252  
 ; Sequence 1252, Application US/09833381  
 ; Patent No. 6672186  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robison, Keith E  
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
 ; FILE REFERENCE: 5800-119  
 ; CURRENT APPLICATION NUMBER: US/09/833,381  
 ; CURRENT FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: 09/516,448  
 ; PRIOR FILING DATE: 2000-02-29  
 ; NUMBER OF SEQ ID NOS: 2050  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1252  
 ; LENGTH: 409  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(409)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-833-381-1252

Alignment Scores:  
 Pred. No.: 1,12e-16 Length: 409  
 Score: 198.50 Matches: 35  
 Percent Similarity: 64.18% Conservative: 8

Best Local Similarity: 52.24% Mismatches: 23  
 Query Match: 10.64% Indels: 1  
 DB: 4 Gaps: 1  
 US-09-466-778B-11 (1-353) x US-09-833-381-1252 (1-409)  
 QY 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21  
 Db 209 ACCGGTCCGACACACCGGCTGTGAATGCCAGTAGGCTACGTGGGTGATGGCTGCAG 268  
 QY 22 CysGluProGlu---GlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 Db 269 TGCTANAGAGCTTGAACCCCTGTGGACAGATGTTGGAGGATCATCACCTTGCCAC 328  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 329 ACTGATGCTTGTGCACTGACCTACATTTCCAGGAAAACAGGCTGGTGTCTTCCACATC 388  
 QY 61 ArgSerProLeuGlyGlnTyr 67  
 Db 389 CAGGCCACCAGTGGCCCTTAT 409  
 RESULT 11  
 US-08-225-477B-1  
 ; Sequence 1, Application US/08225477B  
 ; Patent No. 5633370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Susan Hockfield  
 ; APPLICANT: Diane M. Jaworski  
 ; TITLE OF INVENTION: BEHAV, A Brain Hva-  
 ; TITLE OF INVENTION: luronan-Binding Protein  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: St. Onge Steward Johnston & Reens  
 ; STREET: 986 Bedford Street  
 ; CITY: Stamford  
 ; STATE: CT  
 ; COUNTRY: United States  
 ; ZIP: 06905  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" 1.44 Mb diskette  
 ; COMPUTER: IBM PC  
 ; OPERATING SYSTEM: MS DOS  
 ; SOFTWARE: Word Processor  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/225,477B  
 ; FILING DATE: April 8, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mary M. Krinsky  
 ; REGISTRATION NUMBER: 32423  
 ; REFERENCE/DOCKET NUMBER: 1751-P0004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 203-324-6155  
 ; TELEFAX: 203-327-1096  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1520 bases  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE:  
 ; DESCRIPTION: DNA encoding a protein  
 ; FRAGMENT TYPE: entire sequence  
 ; IMMEDIATE SOURCE: rat brain  
 ; FEATURE:  
 ; NAME/KEY: rat BEHAV  
 US-08-225-477B-1

Alignment Scores:  
 Pred. No.: 2.45e-10 Length: 1520  
 Score: 157.50 Matches: 46  
 Percent Similarity: 41.21% Conservative: 22  
 Best Local Similarity: 27.88% Mismatches: 60



```

;
; TELEPHONE: 203-324-6155
; TELEFAX: 203-327-1096
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1520 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; MOLECULE TYPE:
; DESCRIPTION: DNA encoding a protein
; FRAGMENT TYPE: entire sequence
; IMMEDIATE SOURCE: rat brain
; FEATURE:
; NAME/KEY: rat BEHAB
;
PCT-US95-04353-1

```

Alignment Scores:		Length: 1520	
Pred. No.:	2,45e-10	Matches:	1520
Score:	157.50	Mismatches:	46
Percent Similarity:	41.21%	Conservative:	22
Best Local Similarity:	27.88%	Indels:	37
Query Match:	8.45%	Gaps:	5
DB:			

US-09-466-778B-11 (1-353) x PCT-US95-04353-1 (1-1520)

QY	10	GlucylSerHisTyrValGlyAspGlyLeuAsnCysGluProGluGlnLeuProIle	29
Db	851	CAGTGTGATGGTGGCTGGTGTCCGACCAACCGTAGGTACCCCATCCAGAACCCACGA	910
QY	30	AspArgCysLeuGlnAsp	35
Db	911	GAAGCCTGTTATGGAGACATGGATGCTACCCCTGGATGCGGAATTACGGAGTGGTGGT	970
QY	36	-----AsnGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPhe	50
Db	971	CCTGATGATCTTACGATGTCTACTGTATGCCGAA-----GACCTAAAT---	1015
QY	51	GlnAspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrTyrLeuThr	70
Db	1016	-----GGAGNACTGTTCTAGGTGCCCTCCCGC-----AAGCTGACG	1054
QY	71	PheAspAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGln	90
Db	1055	TGGAGGAGGCTCGGAGTACTGTCTGGAACGGGTGCTCAGATCGCTAGCAGCGGCAG	1114
QY	91	LeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGly	110
Db	1115	CTATACGGCATGGAATGGCGCTTGGACAGATGAGCCCTGGCTGGCTGATGGC	1174
QY	111	ArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIle	130
Db	1175	AGTGTGGGTACCCCATCATCACGCCAGCAACGCTGTGGGGAGGCCCTGCCAGGATC	1234
QY	131	ValAspTyrGlyProArgProAsn-----LysSerGluMetTyrAsp	144
Db	1235	AAGACCTCTTCTCTTTTCCCAACCAAGATGGCTTCCCAAGCAAGAACCGCTTCAAT	1294
QY	145	ValPheCysTyrArg	149
Db	1295	GTCTACTGCTTCGA	1309

RESULT 13

US-08-340-428B-1

Sequence 1, Application US/08340428B

Patent No. 5648465

GENERAL INFORMATION:

APPLICANT: MARGOLIS, Richard U.

APPLICANT: RAUCH, Uwe

APPLICANT: MARGOLIS, Renee K.

TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A

TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE

NUMBER OF SEQUENCES: 49

PROTEOGLYCAN





GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 549 Seconds  
(without alignments)

2731.537 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPGKHCKECKSHYVDGL.....ALAAYSYFRINRKTIGXHF 353

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09466778/runat\_01042004\_095641\_10565/app\_query.fasta\_1.519  
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09466778 @CGN 1.1 470 @runat\_01042004\_095641\_10565 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002s:\*

7: Geneseqn2003as:\*

8: Geneseqn2003bs:\*

9: Geneseqn2003cs:\*

10: Geneseqn2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1849	99.1	1259	3	Aaa57365 DNA encod
2	1772	95.0	4576	6	AbA04648 Rat Hyalu
3	1772	95.0	4642	7	Abx13822 Human Hya
4	1772	95.0	4962	7	Abx13823 cDNA enco
5	1769	94.9	3260	6	Abt08489 Human nov
6	1558	83.5	3625	4	Aaf87120 NOV9 codi
7	1558	83.5	8495	6	Abt08488 Human nov
8	1554	83.3	8444	6	Abt08487 Human nov

9	1514.5	81.2	4706	6	ABA04648	AbA04648 Rat Hyalu
10	1514.5	81.2	4706	7	ABX13821	Abx13821 cDNA enco
11	1275	68.4	1377	3	AAC76373	Aac76373 Human ORF
12	1203	64.5	2011	4	Aaf87113	Aaf87113 NOV2 codi
13	661.5	35.5	3681	6	ABL90758	AbL90758 Human pol
14	658.5	35.3	7879	8	ACF04000	Acf04000 Human CLE
15	656.5	35.2	1522	3	AAA57363	Aaa57363 DNA encod
16	645.5	35.2	6761	3	AAA57362	Aaa57362 DNA encod
17	645.5	34.6	2753	9	ADB47730	Adb47730 Novel hum
18	642	34.4	2160	9	ADB63670	Adb63670 Human CDN
19	632.5	33.9	1482	7	ACA55632	Aca55632 Human sig
20	626	33.6	1804	4	AAF87114	Aaf87114 NOV3 codi
21	581	31.2	2863	9	ADB62456	Adb62456 Human CDN
22	535	28.7	2483	4	AAF87119	Aaf87119 NOV8 codi
23	535	28.7	2512	4	AAF87117	Aaf87117 NOV6 codi
24	363	19.5	474	8	ACH45565	Ach45565 Human foe
25	272.5	14.6	781	7	ABX74436	Abx74436 Human CDN
26	267.5	14.3	1144	7	ABX63446	Abx63446 Human CDN
27	267.5	14.3	1144	9	ADE25720	Ade25720 Human CDN
28	267.5	14.3	1422	7	ACC72678	Acc72678 Human can
29	267.5	14.3	1422	7	ABX76382	Abx76382 Lung canc
30	267.5	14.3	1728	7	ABX63445	Abx63445 Human CDN
31	266.5	14.3	1414	2	AAQ27190	Aaq27190 Tumour ne
32	266.5	14.3	1414	2	AAV71778	Aav71778 Tumour ne
33	266.5	14.3	1414	4	AAD06019	Aad06019 Human tum
34	266.5	14.3	1414	6	ABSS4635	Abss4635 Human CDN
35	256.5	13.8	1411	4	AAH23114	Aah23114 Osteoarth
36	255.5	13.7	1734	6	ABST70406	Abst70406 Human bon
37	248	13.3	351	6	ABV95472	Abv95472 Human pan
38	243.5	13.1	1414	6	ABL68341	AbL68341 Kidney ca
39	243.5	13.1	1414	6	ABK83989	Abk83989 Human CDN
40	243.5	13.1	1414	7	ACC72679	Acc72679 Human can
41	243.5	13.1	1414	7	ABX76383	Abx76383 Lung canc
42	231	12.4	396	8	ACH30555	Ach30555 Human tes
43	213	11.4	718	5	AAS00854	Aas00854 Human CDN
44	202.5	10.9	1365	4	AAI58121	Aai58121 Human pol
45	200.5	10.8	1984	2	AAx52252	Aax52252 Protein p

## ALIGNMENTS

RESULT 1

AAA57365

ID AAA57365 standard; DNA; 1259 BP.

XX

AC AAA57365;

DT 03-OCT-2000 (first entry)

DE DNA encoding a human hyaluronan-binding protein, designated BM-HABP.

XX Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;

KW proliferative condition; metastasis; inflammation; ischemia;

KW host defence dysfunction; immune surveillance dysfunction; arthritis;

KW multiple sclerosis; autoimmunity; immune dysfunction; allergy; se.

XX Homo sapiens.

XX Key Location/Qualifiers

PH CDS

FT 199..1257

FT /tag= a

FT /product= "hyaluronan-binding protein"

FT /transl\_except= (pos: 478..480, aa: Xaa)

FT /transl\_except= (pos: 667..669, aa: Xaa)

FT /transl\_except= (pos: 1105..1107, aa: Xaa)

FT /transl\_except= (pos: 1138..1140, aa: Xaa)

FT /transl\_except= (pos: 1158..1158, aa: Xaa)

FT /transl\_except= (pos: 1168..1170, aa: Xaa)

FT /transl\_except= (pos: 1171..1173, aa: Xaa)

FT /transl\_except= (pos: 1249..1251, aa: Xaa)

FT /note= "Xaa is an unspecified amino acid; no termination codon given"

XX

PN W0200039166-AL.  
 XX PD 06-JUL-2000.  
 XX PF 20-DEC-1999; 99WO-US030462.  
 XX PR 23-DEC-1998; 98US-0113871P.  
 XX FA (HUMA-) HUMAN GENOME SCI INC.  
 XX FA (AMNA-) AMERICAN NAT RED CROSS.  
 XX PI Hastings GA, Liau G, Tsifrina E;  
 XX DR WPI; 2000-452376/39.  
 XX DR P-PSDB; AAY93913.  
 XX PT New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-HABP and BM-HABP, useful for treating proliferative conditions, metastasis, inflammation, ischemia, arthritis and multiple sclerosis.  
 XX PS Claim 2; Fig 4A-B; 457pp; English.  
 CC The present sequence encodes a hyaluronan-binding protein. The specification describes four hyaluronan-binding proteins, known as WF-HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for treating diseases such as proliferative conditions, metastasis, inflammation, ischemia, host defence dysfunction, immune surveillance dysfunction, arthritis, multiple sclerosis, autoimmunity, immune dysfunction and allergy  
 XX SQ Sequence 1259 BP; 295 A; 348 C; 346 G; 258 T; 0 U; 12 Other;

Alignment Scores:  
 Pred. No.: 5,24e-203 Length: 1259  
 Score: 1849.00 Matches: 353  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.14% Indels: 0  
 DB: 3 Gaps: 0

US-09-466-778B-11 (1-353) x AAAS7365 (1-1259)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 DB 199 ATGACAGCCCGGCGACACACAGTGTGAGTAAAGTCACTATGTGGAGATGGGTG 258  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyClnCysHis 40  
 DB 259 AACTGTGAGCCGGAGCAGCTGCCATTGACCGCTGCTTACAGGACAAATGGGCAGTGCCAT 318  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 DB 319 GCAGACGCCAAATGTGTGACCTCCACTCCAGNATACACTGTGGGTGTTCATCTA 378  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 DB 379 CGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGCCCTGTGCCAAC 438  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 DB 439 GAAGCTGGGACCATGGCAACCTTACACAGCTCTCTCTATNNCCAGAAAGGCCAATACAC 498  
 QY 101 LeuCysSerAlaGlyTyrLeuGlnThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 DB 499 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGTGCTTACCCACAGCCCTTGGCCCTCC 558  
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 DB 559 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGACCTATGACCTAGACCCCAACAGAGT 618  
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 DB 619 GAAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAATGCACTNCAAGTGGGC 678

QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180  
 DB 679 TATGTGGAGATGGCTTCTCATACAGTGGGAACCTGTGTGAGGTCTCTGCTCTCC 738  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 DB 739 TCACTCAAAACTTCTGACGGAAGTGTGCTGCTTCCAAACAGCTCAGCTCGAGGCGGT 798  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 DB 799 GCAITTTCTAGAACACCTGACTGACCTGTCCATCGCGGCACCTCTTTTGNCCACAGAAC 858  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 DB 859 AGTGGGTGGGGAGAAATGAGACCTTCTGTGGCGGGACATCGAGCACCACCTCGCCAAT 918  
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 DB 919 GTACAGCATGTTTTCTTCAATGACCTTGTCAATGGCACCCCTGCAACGAGGCTGGGA 978  
 QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 DB 979 AGCAAGCTGTCTCATCTGACAGACAGACCCCACTCCACCCGCGAGACCCAGGTGTGT 1038  
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValle 300  
 DB 1039 GATGGAAGAGACACTCTGGAGTGGGACATCTGTGCTCCCAATGGGATCACCATGTCT 1098  
 QY 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 DB 1099 TCCAGGYCTTTAAAGCACCCCTGCCCCGCTGACCTTGNCCACACTGNTTGGGAGNA 1158  
 QY 321 GlyIlePhe\*\*\*\*\*IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
 DB 1159 GGGATCTTCTGNCATCATCTGCTGACTGGGGCTGTGCTTGGCTTACTCTCTACT 1218  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 DB 1219 TTTCGGATAAACCGGAAACAAATCGGCTTCCANCATTTT 1257

RESULT 2  
 ABA04662  
 ID ABA04662 standard; cDNA; 4576 BP.  
 XX AC ABA04662;  
 XX DT 22-FEB-2002 (first entry)  
 XX DE Human Hyaluronic Acid Receptor for Endocytosis, HARE, coding sequence.  
 XX KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;  
 XX KW chondroitin sulphate; extracellular matrix; cartilage; skin;  
 XX OS vitreous humour; endocytic receptor; glycosaminoglycan; human; ss.  
 XX PH Homo sapiens.  
 XX FT Key Location/Qualifiers  
 XX FT CDS 1..4185  
 XX FT /\*tag= a  
 XX FT /partial  
 XX FT /product= "Human HARE"  
 XX FT /note= "No start codon given"  
 XX PN W0200181544-A2.  
 XX PD 01-NOV-2001.  
 XX PF 25-APR-2001; 2001WO-US013403.  
 XX PR 25-APR-2000; 2000US-0199538P.  
 XX PR 02-NOV-2000; 2000US-0245320P.  
 XX

	QY	181	SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg	200
	Db	3496	TCACTCACAAATCTCTCGAGGAAGTGCTGGCCTATTCCAACAGCTCAGTCGCAGGCCGT	3555
	QY	201	AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn	220
	Db	3556	GCATTTCTAGAACACTGACTGACCTGTCCATCCGGCGCACCCCTCTTTGTGCCACAGAAC	3615
	QY	221	SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn	240
	Db	3616	AGTGGGCTGGGGAGAAATGAGACTTTGTCTGGCGGGACATCGAGCACCACTCGCCAAT	3675
	QY	241	ValSerMetPhePheTyrrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly	260
	Db	3676	GTCAGCATGTTTTTTCACAAATGACCTGTGCANGGCACACCCTCTGAAAAGAGGCTGGGA	3735
	QY	261	SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal	280
	Db	3736	AGCAAGCTGCTCATCACTGCCAGCCAGGACCCCACTCCAACCGGAGACCAAGGTTTGTT	3795
	QY	281	AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle	300
	Db	3796	GATGGAAGAGCCCAATCTGCAGTGGGACATCTTTGCTTCCAATGGATFCATTCATGTCAT	3855
	QY	301	SerArg**LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly***	320
	Db	3856	TCCAGSCCTTTAAAGACACCCCTGCCCCGTGACCTTGACCCACACTGSGCTTGGGAGCA	3915
	QY	321	GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr	340
	Db	3916	GGGATCTCTTTGCCATCATCTCTGGTGACTGGGGCTGTTGCTTGGCTGCTTACTCTAC	3975
	QY	341	PheArgIleAsnArgLysThrIleGlyPhe**HisPhe	353
	Db	3976	TTTCGGATTAACCGGAGACAATCGGCTTCACGACATTTT	4014
		RESULT 3		
		ABX13822		
		ID	ABX13822 standard; cDNA; 4642 BP.	
		XX		
		AC	ABX13822;	
		XX		
		DT	19-FEB-2003... (first entry)	
		XX		
		DE	cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #1.	
		XX		
		KW	Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;	
		KW	chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;	
		KW	tumour; Gene therapy; human; gene; ss.	
		XX		
		OS	Homo sapiens.	
		XX		
		FH	Key	
		CDS	Location/Qualifiers	
			1..4251	
		FT	/tag= a	
		FT	/product= "HARE"	
		FT	/note= "Hyaluronan receptor for endocytosis"	
		FT	/partial	
		FT	/note= "No start codon given"	
		XX		
		PN	WO200286093-A2.	
		PD	31-OCT-2002.	
		XX		
		XX	25-APR-2002; 2002WO-US013209.	
		XX		
		PR	25-APR-2001; 2001US-00842930.	
		PR	25-APR-2001; 2001US-0286468P.	
		XX		
		PA	(WEIG/) WEIGEL P H.	
		PA	(WEIG/) WEIGEL J A.	
		XX		
		PI	Weigel PH, Weigel JA;	

XX WPI; 2003-093126/08.  
 DR P-PSDB; ABG72499.  
 XX  
 PT Targeting compounds e.g. chemotherapeutic agent to cell of subject  
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
 PT epitope.  
 XX  
 XX Example; Fig 9A; 167pp; English.  
 XX  
 CC The invention describes a method of targeting a compound to a cell or  
 CC tissue of an individual expressing a functionally active hyaluronan (HA)  
 CC receptor for endocytosis (HARE) or a cell that does not express  
 CC functionally active HARE. The method involves using HA molecule, a  
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
 CC method is useful for targeting a compound, preferably a chemotherapeutic  
 CC agent or a radioisotope to cell of an individual, especially a human,  
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
 CC method useful for preventing interaction between a cell having at least  
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
 CC its surface. This second method is useful for preventing metastasis by  
 CC preventing interaction between tumour cells having HA, CD or CDS coat and  
 CC non-tumour cells expressing HARE on its surface. The invention also  
 CC describes a method useful for detecting the presence of HA, CD and CDS in  
 CC a biological fluid. This sequence encodes the human 190kDa Hyaluronan  
 CC receptor for endocytosis (HARE)  
 XX  
 SQ Sequence 4642 BP; 1141 A; 1215 C; 1216 G; 1070 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,79e-193 Length: 4642  
 Score: 1772.00 Matches: 334  
 Percent Similarity: 95.18% Conservative: 2  
 Best Local Similarity: 94.62% Mismatches: 17  
 Query Match: 95.01% Indels: 0  
 DB: 7 Gaps: 0

US-09-466-778B-11 (1-353) X ABX13822 (1-4642)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 Db 3022 ATGACAGCCCGCGCAGACCAAGTGTGAGTGAAGAGTCACTATGTGCGAGATGGCTG 3081  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAenGlyGlnCysHis 40  
 Db 3082 AACTGTGAGCGGAGCAGCTGCCAATTGACCGCTGCTTACAGGACAATGGCGCAGTGCAT 3141  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 3142 GCAGACCCCAATGTGCGACCTCCACTTCAGGATACCACTGTGGGGTGTCCATCTA 3201  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 3202 CGCTCCCACTGGCCAGTATAGCTTACACCAAGCCAGAGAGGCTGTGCCAAC 3261  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 Db 3262 GAAGCTCGACCATGGCAACCTACACCAAGCTTCCTATGCCAAGAGGCGCAAGTACCA 3321  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 3322 CTGTGCTCAGCAGCTGGCTGGAGACCGGGGGTGTGCTTACCCACAGCCCTTCCGCTCC 3381  
 QY 121 GlnAsnCysGlySerGlyValValGlyLysLeuValAspTyrGlyProArgProAsnLysSer 140  
 Db 3382 CAGAACTGTGCTCTGTGTGTGTGGGATAGTGGACTATGACCTAGACCCCAACAGAGT 3441  
 QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 Db 3442 GAAATGTGGGATGTCTTCTGCTATCGATGAAGATGTGAACCTGCACCTGCAGAGTGGC 3501

QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180  
 Db 3502 TATGTGGAGATGGCTTCTCATGAGTGGAACTGCTGAGGTCTCTGTATGTCTCTCC 3561  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 Db 3562 TCACCTCAAACTCTCTGACGGAAGTGTGGCTATTCACACAGCTCAGCTCAGGCGCGT 3621  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 Db 3622 GCATTTCTAGAACACCTGACTGACCTGTCCATCGGGGACCCCTCTTTGTGCCACAGAA 3681  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 Db 3682 AGTGGGCTGGGAGAAATGAGACTTGTCTGGGGCGGACATCGAGCACCCACCTCGCAAT 3741  
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 Db 3742 GTCAGCATGTTTTTCTACATGACCTGTCAATGGCACCCCTGCAAAACGAGGCTGGGA 3801  
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 Db 3802 AGCAAGCTGCTCATCTGCGACGAGGACCCACTCCAAACCGACGAGACGAGCTTGT 3861  
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 Db 3862 GATGGAAGAGCCCAATTGCGAGTGGGACATCTTGGCTCCAAATGGGATCATTCATGTCAT 3921  
 QY 301 SerArg\*\*\*LeuLysAlaProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 3922 TCCAGGCCCTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTGGGAGCA 3981  
 QY 321 GlyIlePhe\*\*\*\*\*IleLeuValThrGlyAlaValAlaLeuAlaLysSerTyr 340  
 Db 3982 GGGATCTTCTTCCCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4041  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 4042 TTTCCGATAAACCGGAGACATCGGCTTCCAGCATTTT 4080

RESULT 4  
 ABX13823  
 ID ABX13823 standard; cDNA; 4962 BP.  
 XX  
 AC ABX13823;  
 XX  
 DT 19-FEB-2003 (first entry)  
 XX  
 DE cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #2.  
 XX  
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
 KW tumour; gene therapy; human; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 CDS 1..4962  
 FT /\*tag= a  
 FT /product= "HARE"  
 FT /note= "Hyaluronan receptor for endocytosis"  
 FT /partial  
 FT /note= "No start codon given"  
 XX  
 PN WO200286093-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 25-APR-2002; 2002WO-US013209.  
 XX  
 PR 25-APR-2001; 2001US-00842930.  
 PR 25-APR-2001; 2001US-0286468P.  
 XX

PA (WEIG/) WEIGEL P H.  
 PA (WEIG/) WEIGEL J A.  
 PI Weigel PH, Weigel JA;  
 XX WPI: 2003-093126/08.  
 DR P-PSDB; ABG72514.  
 XX  
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject  
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
 PT epitope.  
 XX  
 XX Example; Fig 9B; 167pp; English.  
 XX  
 CC The invention describes a method of targeting a compound to a cell or  
 CC tissue of an individual expressing a functionally active hyaluronan (HA)  
 CC receptor for endocytosis (HARE) or a cell that does not express  
 CC functionally active HARE. The method involves using HA molecule, a  
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
 CC method is useful for targeting a compound, preferably a chemotherapeutic  
 CC agent or a radioisotope to cell of an individual, especially a human,  
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
 CC method useful for preventing interaction between a cell having at least  
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
 CC its surface. This second method is useful for preventing metastasis by  
 CC preventing interaction between tumour cells having HA, CD or CDS coat and  
 CC non-tumour cells expressing HARE on its surface. The invention also  
 CC describes a method useful for detecting the presence of HA, CD and CDS in  
 CC a biological fluid. This sequence encodes a longer version of the human  
 CC 190kDa Hyaluronan receptor for endocytosis (HARE) shown in ABG72499  
 CC (Encoded by ABX13822)  
 XX  
 SQ Sequence 4962 BP; 1249 A; 1277 C; 1311 G; 1125 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3,08e-193 Length: 4962  
 Score: 1772.00 Matches: 334  
 Percent Similarity: 95.18% Conservative: 2  
 Best Local Similarity: 94.62% Mismatches: 17  
 Query Match: 95.01% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-466-778b-11 (1-353) x ABX13823 (1-4962)  
 QY 1 MetThrglyProGlyLysHisLysCysGlyCysLysSerHisTyrValGlyAspGlyLeu 20  
 DB 3733 ATGACAGCCCGGGCAAGCACAAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 3792  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAnGlyGlnCysHis 40  
 DB 3793 AACTGTGAGCCGGAGCGAGCTGCCCATTTGACCGTGTATACAGGACATATGGGCGAGTGGCAT 3852  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 DB 3853 GCAGACGCCAAATGTGTGACCTCCATCTCCAGATACCACTGTGGGGTGTTCATCTA 3912  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 DB 3913 CGTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCGCTGTGCCAAC 3972  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 DB 3973 GAAGCTCGGACCATGGCAACCTCAACACAGCTCTCTATGCCAGAGGCAAGTATCCAC 4032  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 DB 4033 CTGTGCTCAGCAGGCTGGTGGAGACCGGGGTGCTTACCCACAGCTTCGCGCTCC 4092  
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 DB 4093 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGACTATGACTAGACCCACAGAGT 4152

QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 DB 4153 GAAATGTGGGATGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACCTGCAGGTGGC 4212  
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180  
 DB 4213 TATGTGGAGATGGCTTCTCATGCAGTGGAAACCTGTCTGAGGTCTCTGTATCTCTCCC 4272  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 DB 4273 TCACCTACAACCTCTGACGGAGTGTGCTCTATTCACACAGCTCAGCTCAGAGGCGGT 4332  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 DB 4333 GCATTTCTAGACACACCTGACTGACCTGTCATCCGGGGCACCTCTTTTGTGCCACAGAAC 4392  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 DB 4393 AGTGGGCTGGGAGAAATGAGACCTTGTCTGGGGGGGACATCGAGCACCACTCGCCAAAT 4452  
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 DB 4453 GTCAGCATGTTTCTACATGACCTGTCAATGGCACCACTCCAAACGAGGCTGGGA 4512  
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 DB 4513 AGCAAGCTGCTCATCTGTCAGCGAGGAGCCACTCCAAACGAGGAGACGAGTGTGTT 4572  
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 DB 4573 GATGGAGAGCCCATCTGTCAGTGGGACATCTTGGCTCCCAATGGGATCATTCATGTCATT 4632  
 QY 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 DB 4633 TCCAGGCTTTAAAGCAGACCCCTGCCCCGCTGACCTTGACCCACACCTGGCTGGAGCA 4692  
 QY 321 GlyIlePhe\*\*\*\*\*IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
 DB 4693 GGGATCTCTTTGCCCATCATCTCTGCTGACTGGGGCTGTGGCTGTGCTTACTCTCTAC 4752  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 DB 4753 TTTCCGATAAACCGAGAACAAATCGGCTTCCAGCATTTT 4791  
 RESULT 5  
 ABT08489  
 ID ABT08489 standard; cDNA; 3260 BP.  
 XX  
 AC ABT08489;  
 XX  
 DT 28-NOV-2002 (first entry)  
 XX  
 DE Human novel protein NOV1c coding sequence SEQ ID NO: 210.  
 XX  
 KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;  
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;  
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;  
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;  
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;  
 KW antiparkinsonian; dermatological; antinfertility; cerebroprotective;  
 KW antiaddictive; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT variation replace(887,C)  
 FT variation /\*tag= a  
 FT variation replace(1034,T)  
 FT variation /\*tag= b  
 FT variation replace(1223,T)  
 FT variation /\*tag= c  
 XX



PN W0200259315-A2.  
 XX PD 01-AUG-2002.  
 XX PF 19-DEC-2001; 2001WO-US0500076.  
 XX PR 19-DEC-2000; 2000US-0256619P.  
 XX PR 19-JAN-2001; 2001US-0262959P.  
 XX PR 28-FEB-2001; 2001US-0272408P.  
 XX PR 20-APR-2001; 2001US-0285189P.  
 XX PR 26-JUL-2001; 2001US-0308039P.  
 XX PR 09-AUG-2001; 2001US-0311266P.  
 XX (CURA-) CUPAGEN CORP.  
 PA Shinkets RA, Patturajan M, Vernet CAM, Casman SU, Malyankar U;  
 XX Shenoy S, Spytek KA, Ganggoli E, Miller C, Boldog F, Li L;  
 XX PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;  
 XX PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;  
 XX PI Rothenberg M;  
 XX WPI; 2002-666903/71.  
 DR P-PSDB; ABJ10588.  
 XX  
 PT New isolated NOVX polypeptides and polynucleotides, useful for  
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.  
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease  
 PT or Alzheimer's disease.  
 XX  
 PS Claim 42; Page 24-25; 363pp; English.  
 CC  
 CC The present invention provides the protein and coding sequences of  
 CC several novel human proteins, designated NOVX. These can be used in the  
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-  
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,  
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral  
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia  
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,  
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,  
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,  
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,  
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or  
 CC graft-versus-host disease. The present sequence is a coding sequence of  
 CC the invention  
 XX  
 SQ Sequence 3260 BP; 774 A; 886 C; 850 G; 749 T; 0 U; 1 Other;  
 Alignment Scores:  
 Pred. No.: 3,678-193 Length: 3260  
 Score: 1769.00 Matches: 333  
 Percent Similarity: 95.18% Conservative: 3  
 Best Local Similarity: 94.33% Mismatches: 17  
 Query Match: 94.85% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-466-778B-11 (1-353) x ABT08489 (1-3260)  
 QY 1 MetThrGlyProGlyValHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 DB 1635 ATGACAGCGCCGGGCAAGCAAGGTGAGTGTAAGAGTCACTATGTCCGAGATGGGCTG 1694  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 DB 1695 AACTGTGAGCCGGAGACAGCTGCCATTGACCGCTCTTACAGGACAAATGGGAGTGGCCAT 1754  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 DB 1755 GCAGACGCCAAATGTGTGACCTCCACCTCCAGGATACCACTGTGGGGTGTTCATCTA 1814  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 DB 1815 CGTCCCCACTGGGCGCAGTATAAGTACCTTTGAAAGCCAGAGAGCGCTGTGCCAAC 1874  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 DB 1875 GAAGCTGGACCATGGCAACCTACCAACAGCTCTCTATGCCAGAGCCCAAGTACCAAC 1934  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 DB 1935 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGGGGTGGCTATCCCCACAGCCCTTCGCCTCC 1994  
 QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140  
 DB 1995 CAGAACTGTGGCTCTGGTGTGGTATAGTGGACTATGACCTAGACCCCAACAAGAGT 2054  
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 DB 2055 GAAATGTGGGATGCTCTTCTGCTATCGATGAAGATGTGAATCGACTGCACCTGCAAGTGGGC 2114  
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180  
 DB 2115 TATGTGGAGATGGCTTCTCATGAGTGGGAACTGCTGCGAGTCTCTGATCTCTTCCCC 2174  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 DB 2175 TCACCTCAAACTTCTCTGACGGAAGTGTCTGCTTATTCACACAGCTCAGCTCGAGGCCGT 2234  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerTyrLeuArgGlyThrLeuPheValProGlnAsn 220  
 DB 2235 GCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACTCTTTGTGCCACAGAAC 2294  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyValArgAspIleGluHisLeuAlaAsn 240  
 DB 2295 AGTGGCTGGGGAGATGAGACCTTGTCTGGCGGGACATCGAGACCCACCTCGCCAAAT 2354  
 QY 241 ValSerMetPheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 DB 2355 GTCAGCATGTTTTTCTCAATGACCTTGTCAATGGCACCACCCCTGCAACAGAGGGTGGGA 2414  
 QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 DB 2415 AGCAAGTGTCTCATCTGCGCAGCCAGGACCCACTCCAAACGAGGAGACCAAGTGTGT 2474  
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 DB 2475 GATGGAGAGCCATCTCTGACGTGGGACATCTTTCCCTCCAAATGGGATCAITCATGTCAIT 2534  
 QY 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 DB 2535 TCCAGGCTTTTAAAGACACCCCTGCGCCCTGACCTTGACCCACACTGGCTGGGAGCA 2594  
 QY 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
 DB 2595 GGGATCTTTTGGCATCATCTCGTGGTGGGCTGTGGCTGGGCTGTACTCTCTACTCTAC 2654  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 DB 2655 TTTCCGATAAACCGAGAGACATCGGCTTCCAGCATTTT 2693  
 RESULT 6  
 AAF87120  
 ID AAF87120 standard; DNA; 3625 BP.  
 XX  
 AC AAF87120;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE NOV9 coding sequence.  
 XX  
 KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;  
 KW cardiovascular; casein kinase II phosphorylation site; contraction;  
 KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;  
 KW epidermal growth factor; cell development; apoptosis; cell adhesion;  
 KW growth migration; cell structure; motility; cancer; immune disorder;  
 KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;  
 KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;  
 KW

KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;  
 KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.  
 OS Unidentified.

XX Key Location/Qualifiers  
 FT CDS 387..2024  
 FT /\*tag= a  
 FT /product= "NOV9"

XX WO200136638-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-US0311543.

XX 19-NOV-1999; 99US-0166336P.

XX 29-NOV-1999; 99US-0167785P.

XX 08-MAR-2000; 2000US-0187844P.

XX 16-NOV-2000; 2000US-00715417.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Lichenstein H, Vernet C, Fernandes E;

XX WFI; 2001-648134/74.

XX P-PSDB; AAB83365.

XX Novel human polypeptides and the nucleic acids that encode them useful  
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and  
 PT immune disorders.

XX Claim 9; Page 32-33; 141pp; English.

CC This sequence encodes the NOV9 protein. The invention relates to the NOV1  
 CC -NOV16 proteins, and their coding sequences. The proteins have Cytostatic  
 CC ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular  
 CC activities. The sequences may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate NOVX expression. They  
 CC may be used to treat disorders associated with decreased expression by  
 CC rectifying mutations or deletions in a patient's genome that affect the  
 CC activity of protein by expressing inactive proteins or to supplement the  
 CC patients own production of protein. They are used to produce NOVX  
 CC proteins, by inserting the nucleic acid into a cell and culturing it to  
 CC express the protein. The DNA may be used as DNA probes in assays to  
 CC detect and quantitate the presence of similar DNAs in samples, and which  
 CC patients may need restorative therapy. The NOVX protein may also be used  
 CC as antigens in the production of antibodies (Abs) against NOVX and in  
 CC assays to identify modulators of NOVX expression and activity. The anti-  
 CC NOVX Abs and antagonist are used to down regulate expression and  
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX  
 CC in samples. Disorders that may be prevented, diagnosed and/or treated  
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and  
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of  
 CC serine/threonine kinases, and are used to treat kinase-related disorders  
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).  
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF).  
 CC -like super family and are involved in, e.g. regulation of cell  
 CC development, apoptosis, cell adhesion, growth migration, cell structure  
 CC and motility and protein management, and are used to treat cancers,  
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.  
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to  
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-Qr  
 CC syndrome and marfan syndrome

XX SQ Sequence 3625 BP; 896 A; 971 C; 914 G; 844 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.58e-169 Length: 3625  
 Score: 1558.00 Matches: 300  
 Percent Similarity: 85.31% Conservative: 2  
 Best Local Similarity: 84.75% Mismatches: 15  
 Query Match: 83.54% Indels: 37

DB:	4	Gaps:	1
US-09-466-778B-11 (1-353) x AAP87120 (1-3625)			
Qy	1	MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu	20
Db	2078	ATGACAGGCGCGGCAAGCAAGTGTGAGTGTAAAAAGTCACTATGTCGAGATGGCTG	2137
Qy	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
Db	2138	AACTGTGAGCGGAGGAGCTGCCCAATTGACCGTCTTACAGCAATGGCAGTGCAT	2197
Qy	41	AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu	60
Db	2198	GCAGAGCCCAATGTGTGCGACCTCCACTCCAGGATACCACTGTGGGTGTTCATCTA	2257
Qy	61	ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	2258	CGCTCCCACTGGGCGAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC	2317
Qy	81	GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis	100
Db	2318	GAAGCTGGACCATGGCAACCTACACCACTCTCTATGCCAGAGGCCAAGTACCAC	2377
Qy	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer	120
Db	2378	CTGTGCTCAGAGGCTGGCTGGAGACGGCGGGGTGGCTTACCCACACAGCCTTCGCCCTC	2437
Qy	121	GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	2438	CAGAATGTGGCTCTGTGTGGTGGATAGTAGACTATGGACCTAGACCCACCAAGAGT	2497
Qy	141	GluMetTrpAspValPheCysTyrArgMet - LysAspValAsnCysThr***LysValG1	160
Db	2498	GAATGTGGGATGTCTTCGTATCGGATGAAAG-----	2531
Qy	160	YTrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr	180
Db	2531	-----	2531
Qy	180	oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr	200
Db	2532	-----GAAGTGTGGGCTATTCCAAAGCTCACTCGAGGCGG	2569
Qy	200	GAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs	220
Db	2570	TGCATTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCTCTTTTGTGCCACAGAA	2629
Qy	220	nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaA9	240
Db	2630	CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGACATCGAGCACCACCTCGCAA	2689
Qy	240	nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1	260
Db	2690	TGTCAGCATGTTTTCTACATGACCTTGTCAATGGCACCCCTGCAACAGAGGCTGGG	2749
Qy	260	YSerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa	280
Db	2750	AAGCAAGCTGCTCATCTACCTGCCAGCCAGGACCCCACTCCAAACCGAGAGACCAAGTTGT	2809
Qy	280	lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisVal11	300
Db	2810	TGATGAAGAGGCCATTCTGTCAGTGGGACATCTTTGCTCCAAATGGGATCATTCATGTCAT	2869
Qy	300	eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly**	320
Db	2870	TTCAGGCTTTAAAGACCCCTGCCCGGACCTTGACCCACACCTGGCTGGGAGC	2929
Qy	320	*GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTy	340
Db	2930	AGGGAATCTTTTGGCCATCATCTGCTGAGTGGGGCTGTTGCTTGGCTGCTTACTCTTA	2989
Qy	340	rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe	353

Db 2990 CTTTCGGATAACCGAGAACAAATGGCTTCCAGCATTTT 3029

RESULT 7

ABT08488

ID ABT08488 standard; cDNA; 8495 BP.

AC ABT08488;

DT 28-NOV-2002 (first entry)

XX Human novel protein NOV1b coding sequence SEQ ID NO: 3.

XX Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;  
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;  
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;  
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;  
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;  
 KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;  
 KW antiaddictive; gene; ss.

XX Homo sapiens.

XX WO200259315-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US050076.

XX 19-DEC-2000; 2000US-0256619P.

XX 19-JAN-2001; 2001US-0262959P.

XX 28-FEB-2001; 2001US-0272408P.

XX 20-APR-2001; 2001US-0285189P.

XX 26-JUL-2001; 2001US-0308039P.

XX 09-AUG-2001; 2001US-0311266P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;

PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;

PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;

PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;

PI Rothenberg M;

XX WPI; 2002-666903/71.

XX P-PSDB; ABJ10587.

XX New isolated NOVX polypeptides and polynucleotides, useful for  
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.  
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease  
 PT or Alzheimer's disease.

XX Claim 8; Page 21-23; 363pp; English.

XX The present invention provides the protein and coding sequences of  
 CC several novel human proteins, designated NOVX. These can be used in the  
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-  
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,  
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral  
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia  
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,  
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,  
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,  
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,  
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or  
 CC graft-versus-host disease. The present sequence is a coding sequence of  
 CC the invention

XX SQ Sequence 8495 BP; 2295 A; 2102 C; 2135 G; 1957 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.: 3.39e-168 Length: 8495

Score: 1558.00 Matches: 300

Percent Similarity: 85.31% Conservative: 2  
 Best Local Similarity: 84.75% Mismatches: 15  
 Query Match: 83.54% Indels: 37  
 DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x ABT08488 (1-8495)

QY	1	MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu	20
Db	6948	ATGACAGGCGCGGCAAGCACAAGTGTGAGTAAAGTCACTATGTGCGAGATGGCTG	7007
QY	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
Db	7008	AACGTGTAGCGCGAGCAGCTGCCCATTTGACCGTGTCTTACAGGACATGGCGAGTGC	7067
QY	41	AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu	60
Db	7068	GCAGACGCCAAATGTGTGACCTCCACTTCCAGATACCACTGTGGGGTGTTCATCTA	7127
QY	61	ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	7128	CGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC	7187
QY	81	GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis	100
Db	7188	GAAGCTCGACCATGGCACTCAACACGCTCTCTATGCCAGAAAGGCCAAGTACCAAC	7247
QY	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer	120
Db	7248	CTGTGCTCAGCAGCTGGCTGGAGACCGGGCGGGTTGCCCTACCCACAGCCTTCGCGCTCC	7307
QY	121	GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	7308	CAGAACTGTGGCTCTGGTGTGGATAGTGGACTATGGACCTAGACCAACAGAGT	7367
QY	141	GluMetTrpAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValG1	160
Db	7368	GAATGTGGGATGTCTTCTGCTATCGATGAAG-----	7401
QY	160	yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr	180
Db	7401	-----	7401
QY	180	oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr	200
Db	7402	-----GAAAGTGTGGCTATTCCACAGCTCAGCTCGAGCGCG	7439
QY	200	gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs	220
Db	7440	TGATATTCTAGAACACCTGACTGCTCCATCCCGGGCACCCTCTTTGTGCCACAGAA	7499
QY	220	nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs	240
Db	7500	CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGACATCGAGCACCCCTGCCCAA	7559
QY	240	nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1	260
Db	7560	TGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCCCTGCAACACGAGGTGGG	7619
QY	260	ySerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa	280
Db	7620	AAGCAAGCTGCTCATCTACGCCAGCCAGGACCCACTCCAAACCGAGGACCAAGTTGT	7679
QY	280	lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisVal11	300
Db	7680	TGATGGAAGAGCCATTCTGCAAGTGGGACATCTTTGCTCCATGGGATCATTCATGTCA	7739
QY	300	eSerArg***LeuLysAlaProAlaProValThrLeu***HisThrGlyLeuGly**	320
Db	7740	TTCCAGGCTTTAAAGACACCCCTGCCCGGACCTTGACCCACACTGGCTGGGTGGAGC	7799
QY	320	*GlyIlePhe*****IleIleLeuValThrGlyValAlaValAlaLeuAlaTyrSerTy	340





Db 3859 GATGGAAGATCCATCTGCGAGTGGGACATCATCCGCCCAATGGAATCCCTCCATATTATT 3918  
 QY 301 SerArg\*\*\*LeuLysAlaProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 3919 TCTGAACCTTTGAGAGTCTCTCCACGGCAGCAAGCGCTGCCACTCTGGCTGGGACACA 3978  
 QY 321 GlyIlePhe\*\*\*\*\*IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
 Db 3979 GGTATATTCTGTCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4038  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 4039 TTCGGCTAAAGCAGCGAACCATCTGGTTTCAGCGCTTT 4077  
 RESULT 10  
 ID ABX13821  
 XX ABX13821 standard; cDNA; 4706 BP.  
 AC ABX13821;  
 XX  
 DT 19-FEB-2003 (first entry)  
 XX  
 DE cDNA encoding rat 175kDa Hyaluronan receptor for endocytosis (HARE).  
 XX  
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;  
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;  
 KW tumour; gene therapy; rat; gene; ss.  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4296  
 FT /\*tag= a  
 FT /product= "HARE"  
 FT /note= "Hyaluronan receptor for endocytosis"  
 FT /partial  
 FT /note= "No start codon given"  
 XX  
 PN WO200286093-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 25-APR-2002; 2002WO-US013209.  
 XX  
 PR 25-APR-2001; 2001US-00842930.  
 PR 25-APR-2001; 2001US-0286468P.  
 XX  
 PA (WEIG/) WEIGEL P H.  
 PA (WEIG/) WEIGEL J A.  
 XX  
 PI Weigel PH, Weigel JA;  
 XX  
 DR WPI; 2003-093126/08.  
 DR P-PSDB; ABG72498.  
 XX  
 PT Targeting compounds e.g. chemotherapeutic agent to cell of subject  
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,  
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE  
 PT epitope.  
 XX  
 PS Example; Fig 2; 167pp; English.  
 XX  
 CC The invention describes a method of targeting a compound to a cell or  
 CC tissue of an individual expressing a functionally active hyaluronan (HA)  
 CC receptor for endocytosis (HARE) or a cell that does not express  
 CC functionally active HARE. The method involves using HA molecule, a  
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or  
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The  
 CC method is useful for targeting a compound, preferably a chemotherapeutic  
 CC agent or a radioisotope to cell of an individual, especially a human,  
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a  
 CC method useful for preventing interaction between a cell having at least

CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on  
 CC its surface. This second method is useful for preventing metastasis by  
 CC preventing interaction between tumour cells having HA, CD or CDS coat and  
 CC non-tumour cells expressing HARE on its surface. The invention also  
 CC describes a method useful for detecting the presence of HA, CD and CDS in  
 CC a biological fluid. This sequence encodes the rat 175kDa Hyaluronan  
 CC receptor for endocytosis (HARE)  
 XX  
 SQ Sequence 4706 BP; 1168 A; 1273 C; 1230 G; 1035 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,48e-163 Length: 4706  
 Score: 1514.50 Matches: 280  
 Percent Similarity: 87.54% Conservative: 29  
 Best Local Similarity: 79.32% Mismatches: 43  
 Query Match: 81.21% Indels: 1  
 DB: Gaps: 1  
 US-09-466-778B-11 (1-353) x ABX13821 (1-4706)  
 QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 Db 3022 ATGACGGGCCAGCGCAAGCATAGTGTGAATGTAAAGTCACTATGTGCGGGACGAGTG 3081  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspGlnGlnCysHis 40  
 Db 3082 GACTGTGAGCTGAGCAGCTGCCGCTGACCGTGTCTTACAGGACACAGGACAGTGCAC 3141  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 3142 CCAGATGCCAGCTGTGCAGACCTTACTTCCAGACACAGCCGTAGGAGTATTCATCTA 3201  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 3202 CGTCCCCACTGGCCAGTACAACTGACATTTGACAAAGCCAAAGAGCCTGTGCCAAA 3261  
 QY 81 GluAlaAlaThrMetalThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 Db 3262 GAAGCTCGCACCATAGCCACTCAACACAGCTCTCTATGCCAGAGGCCAAGTATCAC 3321  
 QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyValGlnValAlaTyrProThrAlaPheAlaSer 120  
 Db 3322 CTGTCTCGGCCCGCTGGCTGGAGAGTGGCGGGTGGCTTACCCTACCTACGTATGCTCT 3381  
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 3382 CAGAAGTGTGGTCAAAACGTTTGTGGATCGTAGACTACGGATCCAGGGGCCAACAGAGT 3441  
 QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 Db 3442 GAAATGTGGGATGTCTTCTGTACCGGATGAAAGATGTGAACCTGCACCTGCAAGGAGGC 3501  
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180  
 Db 3502 TATGTGGGAGATGGCTTCTCGTCACTGGGAACCTGCTGCAGGCTCTCATGCTCTCCC 3561  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 Db 3562 TCGCTCACAACTTCCTCTGACAGAGTGTGGCTTTTCCAAAGAGCTCAGCCCGAGGACAG 3621  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 Db 3622 GCATTTTGAACACCTGTGACTGACCTGTCCATCCGTGGCCACCTGTTTGTGCCACAGAAC 3681  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 Db 3682 AGTGGGCTACCGGGAATTAAGACCTGTCTGGCCGGGACATTTAGGACACACCTCATAT 3741  
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 Db 3742 GTCAACGTCTCTCTTTTACAATGACCTGTCAATGGTACCTTTCTGAGGACTATGCTGGGA 3801  
 QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrThrArgCysVal 280

Db 3802 AGCCAACTGCTCATTTACCTTCAGCAGGACACAGCTCCAC---CAAGAGACAGGTTTGTG 3858  
 Qy 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 Db 3859 GATGGAAGATCCATCTTCGAGTGGGACATCATCGCGGCAATGGAATCTCCATATTATT 3918  
 Qy 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 3919 TCTGAACCTTTGAGAGCTCTCCACCGGACGAGCAACGGCTGCCACCTCTGGCTGGGGACA 3978  
 Qy 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaIleThrSerTyr 340  
 Db 3979 GGTATTCTGTGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4038  
 Qy 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 4039 TTCGGCTAAAGCAGCGAACCACTGGTTTCCAGCGCTTTT 4077  
 RESULT 11  
 AAC76373  
 ID AAC76373 standard; cDNA; 1377 BP.  
 AC AAC76373;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF1928 polynucleotide sequence SEQ ID NO:3855.  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnery; antiprosiatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antianemic; antifungal; antirheumatic; antithyroid;  
 KW neurodegenerative disorder; cancer; proliferative disorder; hypertension;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 XX Homo sapiens.  
 OS  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI: 2000-602362/57.  
 DR P-PSDB; AAB42164.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 5; Page 3006-3007; 5507pp; English.  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;

CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antinflammatory; antibacterial;  
 CC antitival; antibacterial; antirheumatic; antithyroid; and antianemic.  
 CC sequences can be used for determining the presence of or predisposition  
 CC to or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 1377 BP; 303 A; 398 C; 342 G; 330 T; 0 U; 4 Other;

## Alignment Scores:

Pred. No.: 1.04e-136 Length: 1377  
 Score: 1275.00 Matches: 252  
 Percent Similarity: 92.73% Conservative: 3  
 Best Local Similarity: 91.64% Mismatches: 17  
 Query Match: 68.36% Indels: 3  
 DB: 3 Gaps: 0

US-09-466-778B-11 (1-353) x AAC76373 (1-1377)

Qy 82 AlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHisLeu 101  
 Db 4 GCTGGACCATGGCAACCTACACAGCTCTCTATGCCCAAGAGGCCAGTACACCTG 63  
 Qy 102 CysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGln 121  
 Db 64 TGCTCAGCAGGCTGGCTGGAGACCGGGGGTGGCTACCCACACAGCCTTCGCTCCACG 123  
 Qy 122 AsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSerGlu 141  
 Db 124 AACTGTGGCTCTGGTGTGGTATGTGACATGTGACCTATGACCTAGACCCCAACAGAGTGA 183  
 Qy 142 MetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGlyTyr 161  
 Db 184 ATGTGGAGATGCTTCTGCTATCGGATGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 243  
 Qy 162 ValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPheProSer 181  
 Db 244 GTGGAGATGGCTTCTCATGCAGTGGGAACCTGTGTGACAGGTCTGTATGCTTCCCTCA 303  
 Qy 182 LeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla 201  
 Db 304 CTCACAACTTCTTGACGGAAGTGTGGCTTATTTCCACAGCTCAGCTCGAGGGCGTGCA 363  
 Qy 202 PheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221  
 Db 364 TTTCTAGAACACCTGACTGACTGTCCATCCGCGGACCCCTCTTTGTGCCACAGACAGT 423  
 Qy 222 GlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsnVal 241  
 Db 424 GGGCTGGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGACACACCTCGCAATGTC 483  
 Qy 242 SerMetPhePheTyrAsnAspLeuValAsnGlyThr-ThrLeuGlnThrArgLeuGlySe 261  
 Db 484 AGCATGTTTTTCTACAATGACCTTGTCAATGGCCACCCNACCTCGAAGAGGGTGGGAAG 543  
 Qy 261 rLysLeuLeuIleThrAspArgGlnAspProLeu-HisProThrGluThrArgCysVal 280  
 Db 544 CAAGCTGCTCATCACTGCCAGCCAGGCCACCCACTNCCNACCCAGCAGGACCGAGTTTGT 603  
 Qy 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 Db 604 GATGGAAGAGCCATTCTGACGTGGGACATCTTTGCTTCCATGGATCATTCATTCATTCATT 663

QY 301 SerArg\*\*\*LeuIysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 664 TCCAGGCGCTTTAAAGCAGCCCTGCGCCGCTGACCTTGACCCACACACTGGCTGGGAGCA 723  
 QY 321 GlyTlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340  
 Db 724 GGGATCTCTTTGGCATCATCTGCTGACCTGGGCTGTTGCTGGCTGCTTACTCTCTAC 783  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 784 TTTCGGATAAACCGAGACACATCGGCTTCAGCATTTT 822  
 RESULT 12  
 ID AAF87113 standard; DNA; 2011 BP.  
 XX AAF87113;  
 XX 26-MAR-2002 (first entry)  
 XX NOV2 coding sequence.  
 XX NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;  
 KW cardiovascular; casein kinase II phosphorylation site; contraception;  
 KW serine/threonine kinase; Feutz-Jeghers syndrome; cellular proliferation;  
 KW epidermal growth factor; cell development; apoptosis; cell adhesion;  
 KW growth migration; cell structure; motility; cancer; immune disorder;  
 KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;  
 KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;  
 KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;  
 KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.  
 XX Unidentified.  
 XX Key Location/Qualifiers  
 FT CDS 877..1824  
 FT /\*tag= a  
 FT /product= "NOV2"  
 XX WO200136638-A2.  
 XX 25-MAY-2001.  
 XX 17-NOV-2000; 2000WO-US031543.  
 XX 19-NOV-1999; 99US-0166336P.  
 XX 29-NOV-1999; 99US-0167785P.  
 XX 08-MAR-2000; 2000US-0187844P.  
 XX 16-NOV-2000; 2000US-00715417.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Lichenstein H, Vernet C, Fernandes E;  
 XX WPI: 2001-648134/74.  
 XX P-PSDB; AAB83358.  
 XX Novel human polypeptides and the nucleic acids that encode them useful  
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and  
 PT immune disorders.  
 XX Claim 9; Page 10-13; 141pp; English.  
 XX This sequence encodes the NOV2 protein. The invention relates to the NOV1  
 CC -NOV16 proteins, and their coding sequences. The proteins have Cytostatic  
 CC ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular  
 CC activities. The sequences may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate NOVX expression. They  
 CC may be used to treat disorders associated with decreased expression by  
 CC rectifying mutations or deletions in a patient's genome that affect the  
 CC activity of protein by expressing inactive proteins or to supplement the  
 CC patients own production of protein. They are used to produce NOVX

CC proteins, by inserting the nucleic acid into a cell and culturing it to  
 CC express the protein. The DNA may be used as DNA probes in assays to  
 CC detect and quantitate the presence of similar DNAs in samples, and which  
 CC patients may need restorative therapy. The NOVX protein may also be used  
 CC as antigens in the production of antibodies (Abs) against NOVX and in  
 CC assays to identify modulators of NOVX expression and activity. The anti-  
 CC NOVX Abs and antagonist are used to down regulate expression and  
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX  
 CC in samples. Disorders that may be prevented, diagnosed and/or treated  
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and  
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of  
 CC serine/threonine kinases, and are used to treat kinase-related disorders  
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).  
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF).  
 CC -like super family and are involved in, e.g. regulation of cell  
 CC development, apoptosis, cell adhesion, growth migration, cell structure  
 CC and motility and protein management, and are used to treat cancers,  
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.  
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to  
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT  
 CC syndrome and marfan syndrome  
 XX  
 XX Sequence 2011 BP; 458 A; 551 C; 583 G; 419 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 3,7e-128 Length: 2011  
 Score: 1203.00 Matches: 230  
 Percent Similarity: 84.87% Conservative: 0  
 Best Local Similarity: 84.87% Mismatches: 4  
 Query Match: 64.50% Indels: 37  
 DB: 4 Gaps: 1  
 US-09-466-778B-11 (1-353) x AAF87113 (1-2011)  
 QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 Db 1309 ATGACAGCCCGCGGCAAGCACAAGTGTGAGTGTAAAGTCACTATGTGCGAGATGGGCTG 1368  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 Db 1369 AACTGTGAGCGCGAGCAGCTGCCATTGACCGCTGTACAGGACAATGGCGAGTGCAT 1428  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 1429 GCAGAGCCCAATGTCCGACCTCCACTTCCAGATACCACTGTGGGGTTCATCTA 1488  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 1489 CGTCCCTCCACTGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 1548  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 Db 1549 GAAAGTCGAGACCATGGCACTACACAGCTCTCTATGCCAGAGGCGCAAGTACCAAC 1608  
 QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 1609 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTGCCTTACCCACAGGCTTCGCGCTC 1668  
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 1669 CAGACTGTGGCTCTGGTGTGGTGGATAGTAGTATGAGCCAGAGGCTGTGCCAAGAGT 1728  
 QY 141 GluMetTrpAspValPheCysTyrArgMet - LysAspValAsnCysThr\*\*\*LysValGI 160  
 Db 1729 GAAATGTGGATGCTCTTCTGCTATCGATGAAAG----- 1762  
 QY 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuValLeuMetSerPhePr 180  
 Db 1762 ----- 1762  
 QY 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200  
 Db 1763 -----GAAATGTGGCTTATTCACAGCTCAGCTCAGGCGCG 1800



Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220  
 Db 1801 TGCATTTCTAGAACACCTGACTGACCTGCCATCGCGGCACCTCTTTGTGCCACAGAA 1860  
 Qy 220 nSerGlyLeuGlyGluAenGluThrLeuSerClyArgAspIleGluHisLeuAlaAs 240  
 Db 1861 CAGTGGGCTGGGGAGAAATGAGACCTTCTGTGGCGGACATCGAACACCTCGCCAA 1920  
 Qy 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG 260  
 Db 1921 TGTGACGATGTTTTCTACAATGACCTTGTCATGTCACACCTCGTCAACAGAGGCTGGG 1980  
 Qy 260 ySerIysLeuLeulleThrAspArgGlnAsp 270  
 Db 1981 AAGCAAGTGTCTCATCACTGCCAGCAGGAC 2011  
 RESULT 13  
 ABL90758  
 ID ABL90758 standard; cDNA; 3681 BP.  
 XX  
 AC ABL90758;  
 XX  
 DT 24-MAY-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 1320.  
 XX  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200190304-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-US016450.  
 XX  
 PR 19-MAY-2000; 2000US-0205515P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI; 2002-122018/16.  
 XX  
 DR P-PSDB; ABB90349.  
 XX  
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 PT prevention of neural, immune system, muscular, reproductive,  
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 PT disorders.  
 XX  
 PS Claim 4; SEQ ID NO 1320; 208lpp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins  
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3681 BP; 648 A; 1172 C; 1123 G; 731 T; 0 U; 7 Other;  
 Alignment Scores:  
 Pred. No.: 2,79e-65 Length: 3681  
 Score: 661.50 Matches: 145  
 Percent Similarity: 53.44% Conservative: 49  
 Best Local Similarity: 39.94% Mismatches: 148  
 Query Match: 35.47% Indels: 21  
 DB: 6 Gaps: 7  
 US-09-466-778B-11 (1-353) x ABL90758 (1-3681)  
 Qy 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21  
 Db 2235 ACCGGCTGAACACACCGCGCTGTGAGTGCACCGAGGCTACGTAGCGATGAGTGCAG 2294  
 Qy 22 Cys--GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 Db 2295 TGTCTGGAGGAGTGGGAACACACCTGTGGACCGCTTGGGCCAGCCACCGCCCTGCCAC 2354  
 Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 2355 TCAGATGCCATGTGCACTGACCTGCATTCAGGAGAAACGGGCTGGGTTTTCACCTC 2414  
 Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 2415 CAGGCCACACGCGCCCTTATGTGCTGAACCTTTTCGAGGCTGAGGCGGATGCGAAGCA 2474  
 Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*GlnLysAlaLysTyrHis 100  
 Db 2475 CAGGAGCGCTGCTTGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2534  
 Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyValAlaTyrProThrAlaPheAlaSer 120  
 Db 2535 CTGTGCTCATGGCTGGCTGGCCAAATGGCTCCACTGCCACCTGTGGTTTTCCTGTG 2594  
 Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 2595 GCGGACTGTGCAATGTGCGGTGGGCTAGTACAGCTGGGTGGCCGCAAGAACCTCTCA 2654  
 Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*LysValGly 160  
 Db 2655 GAACGTGGGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2714  
 Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuGlnValLeuMetSerPhe 179  
 Db 2715 TTCGTGGGTGAGCGGATCAGCAGCTGCAATGGGAAGCTGTGGATGTGTGGCTGCCACT 2774  
 Qy 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199  
 Db 2775 GCCAACTCTCCACCTTCTATGGGATGCTATTGGGCTATGCCATGCCACCCAGCGGGT 2834  
 Qy 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219  
 Db 2835 CTCGACTCTCCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2894  
 Qy 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239  
 Db 2895 AATGAAGCTTTGTGGCAACATGACGCTGAGTGGCCAGACTTGGAGTGCATGCTCC 2954  
 Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259  
 Db 2955 AACGCCACCTCTCTAAGTGGCAAC---GCCAGCCAGGGGAAGTTGCTTCCGCCCATCA 3011  
 Qy 260 GlySerLysLeuLeuIleThrAspArg-----GlnAspProLeuHisProThrGlu 276  
 Db 3012 GGCCTGAGCTTCATCATCAGTGCAGCGAGGCCCTGACAACTGCTTCTGGGCCCTTGTGCC 3071  
 Qy 277 ThrArgCysValAspGlyArgAspThrLeuGluThrPheAspIleCysAlaSerAsnGlyLe 296  
 Db 3072 CCAGGGACAGTTGGTGTAGCGGTATCATTTGTGTGGGACATCATGCGCTTCAATGGCATC 3131

QY 297 ThrHisValIleSerArg\*\*\*LeuLysAlaProProlaProValThrLeu----- 313  
 Db 3132 ATCCATGCTCTGGCAGCCCTCTCTGGCAGCCACCCAGCCAGGAGTGGTGGCGCT 3191  
 QY 314 -----\*\*\*HisThrGlyLeuGly\*\*\*GlyIlePhe\*\*\*\*\*IleIleLeuVal 329  
 Db 3192 GAAGCCCACTGTGGCGGCGGAGGCGTGGGGCT-----GTGCTGGC 3233  
 QY 330 ThrGlyAlaVal-----AlaLeuAlaLysSerTyrPheArgIleAsnArgLysThr 347  
 Db 3234 CTGGAGACATGCTGTGGTGGCGGAGCTCTACCTCCGTGCCGAGGAGCAAGCC 3293  
 QY 348 IleGlyPhe 350  
 Db 3294 ATGGGCTTT 3302

RESULT 14  
 ACF04000  
 ID ACF04000 standard; cDNA; 7879 BP.  
 XX  
 AC ACF04000;  
 DT 09-OCT-2003 (first entry)  
 DE Human CLEVER-1 encoding cDNA SEQ ID NO:1.  
 XX  
 KW Human; lymphatic endothelial glycoprotein; CLEVER-1; inflammation;  
 KW common lymphatic endothelial and vascular endothelial receptor 1;  
 KW antiinflammatory; dermatological; antidiabetic; antirheumatic; lupus;  
 KW antiarthritic; osteopathic; antiasthmatic; cytostatic; hepatotropic;  
 KW antiatherosclerotic; neutropenic; immunosuppressive; skin inflammation;  
 KW acute inflammatory disease; chronic inflammatory disease; diabetes;  
 KW connective tissue disease; rheumatoid arthritis; osteoarthritis;  
 KW lung disease; asthma; adult respiratory distress syndrome; sarcoidosis;  
 KW idiopathic pulmonary fibrosis; inflammatory bowel disease; nephritic;  
 KW ulcerative colitis; Crohn's disease; non-viral hepatitis; cirrhosis;  
 KW cholangitis; atherosclerosis; vasculitis; thyroiditis; multiple sclerosis;  
 KW myositis; ischaemia reperfusion injury; transplantation rejection;  
 KW metastasis; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 31..7743  
 FT FT /\*tag= a  
 FT FT /product= "CLEVER-1"  
 FT FT /note= "common lymphatic endothelial and vascular  
 FT FT endothelial receptor 1"  
 XX  
 PN W02003057130-A2.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 08-JAN-2003; 2003WO-FI000010.  
 XX  
 PR 09-JAN-2002; 2002US-0346298P.  
 XX  
 PA (JALK/) JALKANEN S.  
 PA (IRJA/) IRJALA H.  
 PA (SALM/) SALMI M.  
 XX  
 FI Jalkanen S, Irjala H, Salmi M;  
 DR WPI; 2003-587066/55.  
 DR P-PSDB; ABR82200.  
 XX  
 PT New purified lymphatic endothelial glycoprotein, designated as common  
 PT lymphatic endothelial and vascular endothelial receptor- (CLEVER-1),  
 PT recognizable by a monoclonal antibody, useful for treating inflammation.  
 XX  
 PS Claim 2; Fig 9; 68pp; English.  
 XX

CC The present sequence encodes a human lymphatic endothelial glycoprotein,  
 CC designated as common lymphatic endothelial and vascular endothelial  
 CC receptor 1 (CLEVER-1), which has a molecular weight of 270-300 kD in  
 CC sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)  
 CC under non-reducing conditions, recognisable by a monoclonal antibody  
 CC selected from the group consisting of DSM ACC 2519 or 2590. Also  
 CC described: (1) a CLEVER-1 antibody; (2) a cell-free preparation  
 CC comprising CLEVER-1; (3) diagnosing inflammatory diseases in a patient;  
 CC (4) detecting malignant cells in a patient; (5) identifying an agent that  
 CC stimulates or inhibits the binding of CLEVER-1 to cells; (6) removing  
 CC malignant cells from a sample; (7) treating inflammation in a patient;  
 CC (8) preventing metastasis in a patient; and (9) stimulating CLEVER-1  
 CC binding in a patient. CLEVER-1 has antiinflammatory, dermatological,  
 CC antidiabetic, antirheumatic, antiasthmatic, osteopathic, antiasthmatic,  
 CC cytostatic, hepatotropic, antiatherosclerotic, neutropenic and  
 CC immunosuppressive activities. The CLEVER-1, antibodies and methods from  
 CC the present invention can be used for treating inflammation, e.g. acute  
 CC or chronic inflammatory diseases such as skin inflammations, diabetes,  
 CC connective tissue diseases (such as lupus, rheumatoid arthritis,  
 CC osteoarthritis), obstructive and restrictive lung diseases (such as  
 CC asthma, adult respiratory distress syndrome (ARDS), sarcoidosis,  
 CC idiopathic pulmonary fibrosis), inflammatory bowel diseases (such as  
 CC ulcerative colitis and Crohn's disease), various nephritides, non-viral  
 CC hepatitis, cirrhosis, cholangitis, atherosclerosis, vasculitis,  
 CC thyroiditis, multiple sclerosis, myositis, ischaemia reperfusion injury  
 CC or transplantation rejection, and for preventing metastasis  
 XX  
 SQ Sequence 7879 BP; 1421 A; 2487 C; 2447 G; 1524 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,91e-64 Length: 7879  
 Score: 658.50 Matches: 144  
 Percent Similarity: 53.17% Conservative: 49  
 Best Local Similarity: 39.67% Mismatches: 149  
 Query Match: 35.31% Indels: 21  
 DB: 8 Gaps: 7

US-09-466-778B-11 (1-353) x ACF04000 (1-7879)

QY 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21  
 Db 6487 ACCGGCTGACACACACGGGCTGTGTAGTCCAGCGAGGTACGTAGCGATGACATGCAG 6546  
 QY 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 Db 6547 TGTCTGGAGAGTGGGAGACACACCTGTGTGACCGCTGTGGCCAGCACCGCCCTGCCAC 6606  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 6607 TCAGATGCCATGTGCACTGACCTGCACCTCCAGGAGAAACGGGCTGGCGCTTTCCACCTC 6666  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgLysAlaCysAlaAsn 80  
 Db 6667 CAGGCCACAGCGGCCCTTATGCTGTGAATTTTCGAGGCTGAGCGGATCGGAAGCA 6726  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 Db 6727 CAGGAGACCGTCTCTGCTTCATTCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCAC 6786  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 6787 CTGTGCTCATGGCTGGCTGGCCAAATGGCTCCACTGCCACCCCTGTGTTTCCCTGTG 6846  
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 6847 GCGGACTGTGGCAATGTTCGGGTGGCGGTAGTCAGCTGGGTGCCCGCAAGACCTCTCA 6906  
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 Db 6907 GAACGCTGGGATGCCCTACTCTCCGTGTGCAAGATGTGGCTGTGCCGATGCCGAAATGGC 6966  
 QY 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuValLeuMetSerPhe 179  
 Db 179

```
Db 6967 TTCGTGGGTGACGGGATCAGCACGTGCAATGGGAAGCTGTGGATGTCTGGTGGCCACT 7026
Qy 180 ProSerLeuThrAsnPhLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199
Db 7027 GCCAACTCTCCACTCTTATGGGATGCTATGGCTATGCCATGCCACCGGGGT 7086
Qy 200 ArgAlaPhLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219
Db 7087 CTCGACTTCTGGACTTCTCGATGATGAGCTACGTATAAGACACTCTTGTCTGCTGTC 7146
Qy 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
Db 7147 AATGAAGCTTGTGGACATACGCTGATGGTCCAGACTTGGAGCTGATGCCCTCC 7206
Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
Db 7207 AACGCCACCTCTAAGTGGCAAC---GCCAGCCAGGGGAAGTGTCTCCGGCCACTCA 7263
Qy 260 GlySerIysLeuLeuThrAspArg-----GlnAspProLeuHisProThrGlu 276
Db 7264 GGCCTCAGCCTCATCATCAGTACGACGAGCCCTGACACAGTTCCTGGGCCCTGTGGCC 7323
Qy 277 ThrArgCysValAspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIle 296
Db 7324 CCAGGACAGTGTGGTAGCGGTATCATGTGTGGACATCATGCGCTTCAATGGCATC 7383
Qy 297 ThrHisValIleSerArg***LeuIysAlaProProAlaProValThrLeu----- 313
Db 7384 ATCCATGTCTGGCAGCCCTCTCTGGCACCCCCACAGCCAGGAGTGTGGCGCT 7443
Qy 314 -----***HisThrGlyLeuGly***GlyIlePhe***IleIleLeuVal 329
Db 7444 GAAGCCCACTGTGGCGGAGCGGTGGGGCT-----GTGCTTGCC 7485
Qy 330 ThrGlyAlaVal-----AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgIysThr 347
Db 7486 GCTGGAGCACTGTGGTGTGGTGGCGGAGCTCTTACCTCCGTGCCCGAGGAGGCC 7545
Qy 348 IleGlyPhe 350
Db 7546 ACGGGCTTT 7554

RESULT 15
ID AAA57363 standard; DNA; 1522 BP.
AC AAA57363;
XX
DT 03-OCT-2000 (first entry)
XX
DE DNA encoding a human hyaluronan-binding protein, designated WF-HABP.
XX
KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;
KW proliferative condition; metastasis; inflammation; ischemia;
KW host defence dysfunction; immune surveillance dysfunction; arthritis;
KW multiple sclerosis; autoimmunity; immune dysfunction; allergy; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
XX
FT Location/Qualifiers
XX 31..1404
XX /tag= a
XX /product= "hyaluronan-binding protein"
XX /transl_except= (pos: 1102..1104, aa: Xaa)
XX /transl_except= (pos: 1255..1257, aa: Xaa)
XX /transl_except= (pos: 1297..1299, aa: Xaa)
XX /note= "Xaa is an unspecified amino acid"
XX
FN WO200039166-A1.
XX
PD 06-JUL-2000.
XX
PF 20-DEC-1999; 99WO-US030462.
```

```
XX 23-DEC-1998; 98US-0113871P.
FR (HUMA-) HUMAN GENOME SCI INC.
PA (AMNA-) AMERICAN NAT RED CROSS.
XX
PI Hastings GA, Liao G, Tsifrina E;
XX
XX WPI; 2000-452376/39.
DR P-ESDB; AAY93911.
XX
XX New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-
PT HABP and BM-HABP, useful for treating proliferative conditions,
PT metastasis, inflammation, ischemia, arthritis and multiple sclerosis.
XX
XX Claim 2; Fig 2A-B; 457pp; English.
XX
XX The present sequence encodes a hyaluronan-binding protein. The
CC specification describes four hyaluronan-binding protein, known as WF-
CC HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for
CC treating diseases such as proliferative conditions, metastasis,
CC inflammation, ischemia, host defence dysfunction, immune surveillance
CC dysfunction, arthritis, multiple sclerosis, autoimmunity, immune
CC dysfunction and allergy
XX
SQ Sequence 1522 BP; 277 A; 468 C; 454 G; 312 T; 0 U; 11 Other;
```

## Alignment Scores:

Pred. No.:	2,84e-65	Length:	1522
Score:	656.50	Matches:	144
Percent Similarity:	53.44%	Conservative:	50
Best Local Similarity:	39.67%	Mismatches:	148
Query Match:	35.20%	Indels:	21
DB:	3	Gaps:	7

US-09-466-778B-11 (1-353) x AAA57363 (1-1522)

```
Qy 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
Db 148 ACCGGCCTCAACACACACGCGCTGTGAGTGCACAGGCTACGTAGGCGATGGACTGCAG 207
Qy 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 208 TGTCTGGAGAGTCCGAACACACCTGTGTGACCGCTGTGGCCAGCCAGCCCTGCCAC 267
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 268 TCAGATGCCATGTGCACTGACCTGCACTTCAGGAGAAACGGGCTGGCGTTTCCACCTC 327
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 328 CAGGCCACCGAGCGCCCTTATGGTCTGAACCTTTTCGAGGCTAGGCGGATGCGAAGCA 387
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 388 CAGGAGCGCGCTCTGTGCTTATCCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCAC 447
Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 448 CTGTGCTCATGGGCTGGCTGGCCATATGGTCCACTGCCACCTGCTGTTTCCCTGTG 507
Qy 121 GlnAsnCysGlySerGlyValValGlyValValAspTyrGlyProArgProAsnLysSer 140
Db 508 CGGAGCTGTGGCAATGTGGTGGGATAGTAGCTAGCTGGGTGCCCAAGAACCTCTCA 567
Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 568 GAACGCTGGGATGCTACTCTCTCCGTGTGCAAGATGTGGCTGCCATGCCGAAATGGC 627
Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuGlnValLeuMetSerPhe 179
Db 628 TTCGTGGTGCAGCGGATCAGCACGTGCAATGGGAAGCTGTGGATGTGTGCTGCTGCCACT 687
```

Search completed: April 1, 2004, 12:37:58  
Job time : 586 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 4925 Seconds  
(without alignments)  
3106.617 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865  
Sequence: 1 MTGPKHCKCKSHYVDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09466778/runat\_01042004\_095642\_10579/app\_query.fasta\_1.519  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09466778@cgn 1 3731 @runat\_01042004\_095642\_10579 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pin.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1772	95.0	4492	9	AY227444 Homo sapi
2	1772	95.0	4575	9	AK024503 Homo sapi
3	1772	95.0	5604	9	AK074051 Homo sapi
4	1772	95.0	8214	9	AB052958 Homo sapi
5	1772	95.0	8251	9	AY311388 Homo sapi
6	1772	95.0	8266	9	AJ295695 Homo sapi
7	1769	94.9	3260	6	AX686401 Sequence
8	1769	94.9	3260	6	AX686401 Sequence
9	1558	83.5	3625	6	AX149461 Sequence
10	1558	83.5	3642	9	AL133021 Homo sapi
11	1558	83.5	8495	6	AX686192 Sequence
12	1554	83.3	8444	6	AX686192 Sequence
13	1514	81.2	4706	10	AY007370 Rattus no
14	1487	79.7	8157	10	AF364951 Mus muscu
15	1203	64.5	2011	6	AX149447 Sequence
16	664.5	35.6	3911	10	BC031166 Mus muscu
17	664.5	35.6	7931	10	AF290914 Mus muscu
18	661.5	35.5	7909	9	D87433 Homo sapien
19	661.5	35.5	7916	9	AB052956 Homo sapi
20	652.5	35.0	7870	9	HS275213 Homo sapi
21	643	34.5	1500	4	BTAS50060 Bos tauru
22	642	34.4	2160	6	AX748299 Sequence
23	642	34.4	2160	6	AX093774 Homo sapi
24	632.5	33.9	1482	6	AR269667 Sequence
25	626	33.6	1804	6	AX149449 Sequence
26	581	31.2	2863	6	AX747085 Sequence
27	581	31.2	2863	9	AK091600 Homo sapi
28	535	28.7	2483	6	AX149459 Sequence
29	535	28.7	2512	6	AX149455 Sequence
30	479.5	25.7	249541	2	AC095672 Rattus no
31	479.5	25.7	249774	2	AC126196 Rattus no
32	464	24.9	199310	9	AC112215 Homo sapi
33	464	24.9	217029	9	AC105934 Homo sapi
34	458.5	24.6	146283	2	AC022277 Homo sapi
35	458.5	24.6	180064	9	AC025265 Homo sapi
36	458.5	24.6	180525	2	AC063946 Homo sapi
37	458.5	24.6	184435	2	AC141618 Homo sapi
38	458.5	24.6	238220	2	AC107727 Mus muscu
39	414	22.2	1249	10	BC022136 Mus muscu
40	400.5	21.5	202968	2	AC137422 Rattus no
41	388.5	20.8	204736	2	AC113310 Mus muscu
42	388.5	20.8	247224	2	AC112790 Mus muscu
43	388.5	20.8	343587	2	AC113471 Mus muscu
44	388.5	18.7	6824	6	AX781121 Sequence
45	267.5	14.3	769	9	HUMZD93D04 Homo sapi

# ALIGNMENTS

RESULT 1

AY227444 4492 bp mRNA linear PRI 23-APR-2003  
LOCUS Homo sapiens hyaluronan receptor for endocytosis precursor (HARE)  
DEFINITION mRNA, partial cds.  
ACCESSION AY227444  
VERSION AY227444.1 GI:28848629  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 4492)  
AUTHORS Zhou, B., McGary, C.T., Weigel, J.A., Saxena, A. and Weigel, P.H.  
TITLE Purification and molecular identification of the human hyaluronan  
receptor for endocytosis  
JOURNAL Glycobiology 13 (5), 339-349 (2003)  
MEDLINE 22588541  
PUBMED 12626425  
REFERENCE 2 (bases 1 to 4492)  
AUTHORS Weigel, P.H.  
TITLE Direct Submission  
JOURNAL Submitted (29-JAN-2003) Biochemistry and Molecular Biology,  
University of Oklahoma Health Sciences Center, BMSB Room 860,  
Oklahoma City, OK 73104, USA  
LOCATION/Qualifiers  
FEATURES  
source  
1..4492  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
<1..4492  
/gene="HARE"  
<1..4251  
/gene="HARE"  
/codon\_start=1  
/product="hyaluronan receptor for endocytosis precursor"  
/protein\_id="AAO39681.1"  
/db\_xref="GI:28848630"  
/translation="SLPNLLMRLEQMPDYSIFRGVYIIQYNLANAIEAADAYTVFAPNN  
NAIENYIREKVLSEEDVLVHYVLEKLLKNDLHNGMHRETMLGFSYRFLSLFND  
QLYNEAPINTVATDKGVJHGLKVLKQNRCDNDNTIILGRCTCSSELTCTPP  
TVCSLGNRRKRIYTSYFMGRRTLFIQCPKCRVITRECCAGFFGQCPQFNAG  
NVCFNGICLVGVNGTQCEGEGFSGTACETGKGIHCDQACSVHRCNGQPL  
GDGSCDVGWRGVCDNATNEDNGTCHTANCLTNSDGTASCACAGFQNGGTC  
TAINACEISNGGCSAKAKRTTPGRVCTCKAGTGDIVCLEINPLENHGCDKN  
AECTGTGPNQACNCLPAYTGDKVCTLIYVCLTKNGGCSFALCNHTGQVETCTCK  
PNYIGDFTCRGSIYQELPKPKTSQYFQLOEHEFVLDLVPGPPTVPAPLSAADER  
ARVWDWKYGLMPQVLRHYVACHQLLENLKLISNATSLQGEFIVSVSSTYINN  
KAKLISDIISTNGIVHIIIDKLLSPNLLITPKNSGRILQNLITLATNNGITIKFSNL  
TODGLLSVIITDPIHTVTLFWPTDQALHALPABQQDFLNQDNKDLKEYLKFVIR  
DAKVLADPLTSATWKTQGSSELSVKCGAGRDIGDLFLNGQTCTRIVORELLFDLGVAY  
GIDCLLIDPLTGGRCDTFTFPDASGEGSCVNTSPCRWSPKGVKQKLYNLFPKRN  
LEGCRERSLVITQIPRCCKGVFRDCAQCPGPDAPCNNGVGLDQXATGECKNTG  
PNTGACEMCPGEPGDCPLPGCSDHGCQDGLTGSGQCLCTGTGTPSCDTQAVLPA  
VCTPPCSAHATCKENCTECNIDYEGDITCTVVDVFCQDNGGCAKVARCSQKGTKVS  
CSQCKYKGDHSCSTEDIPCADLNGGCHHEATKMTGPGHKCEKSHYVYGDGANCE  
PEQPIFDLRQDCHADAKCVLDLHQDTTGVFPLRSLQYKLFIDKAREACANE  
AATMATYNQLSVQAKVHLCSAGWLETGRVAYPTAFASQNCGGVGVIVDYGPRPNK  
SEMDVFCYRMKDVNCTCKGVGDFGSCGNLLQVLMRPSLNFELTEVLAYNSSA  
RGRAPLEHLTDLISRGTLFPVQNSGLGENETLSGDIIEHLIANYSMFYNDLVNGLTIL  
QTRIGSKULLITASQDPLQPTETRFVDGRAILQWIDFASNGIIVHSIRPLKAPPAPVIL  
THTGLGAGIFAILVLTGVALAAYSYFRINRRTIGFQHFSEEDINVAALGKQOPEN  
ISNLYESTTSAPPEPSYDPFTDSEERQLEGNPLRTL"  
1..4248  
mat\_peptide  
/gene="HARE"  
/product="190 kDa hyaluronan receptor for endocytosis"  
/function="mediates endocytosis of hyaluronan or  
chondroitin sulfate via the clathrin coated pit pathway"  
/note="generated by proteolytic processing of a larger  
protein; probably a product of the HARE/Stablin-2 gene;  
190 kDa protein purified from spleen"

ORIGIN

Alignment Scores:



```

Db      3795  GATGGAAGAGCAATCTCGACGTGGGACATCTTTGGCTCCAAATGGGATCAATTCATGTCATT 3854
Qy      301  SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db      3855  TCCAGGCGCTTTAAAGACACCCCTCCCGCGTGACCTTTGACCCACACCTGGCTGGGAGCA 3914
Qy      321  GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
Db      3915  GGGATCTCTTTGGCATCATCTCGTGACGTGGGCTGTTCCTTGGCTGCTTACTCTCTAC 3974
Qy      341  PheArgIleAsnAtgLyThrIleGlyPhe***HisPhe 353
Db      3975  TTTGGATAAACCGAGACATCGCTTCCAGCAATTT 4013

RESULT 3
AK074051
LOCUS      AK074051
DEFINITION Homo sapiens mRNA for FLJ00122 protein.
ACCESSION AK074051
VERSION    AK074051.1 GI:18676449
KEYWORDS  fib (full insert sequence).
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
            The nucleotide sequence of a long cDNA clone isolated from human
            spleen
            Published Only in Database (2002)
            2 (bases 1 to 5604)
            Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
            Direct Submission
            Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research
            Institute, Department of Human Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert and 5'- &
            3'-end one pass sequencing; Research Association for Biotechnology;
            cDNA library construction and clone selection; Kazusa DNA Research
            Institute.
FEATURES             Location/Qualifiers
     source           1..5604
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FLJ00122"
                     /tissue_type="spleen"
                     /dev_stages="adult"
                     /note="vector:pBluescriptII SK plus"
     gene             1..5604
                     /gene="FLJ00122"
     CDS              <1..5213
                     /gene="FLJ00122"
                     /note="Start codon is not identified."
                     /codon_start=3
                     /evidence=not experimental
                     /product="FLJ00122 protein"
                     /protein_id="BAB04877.1"
                     /db_xref="GI:18676450"
                     /translation="DGSAGRLCDKQTSACGPYVQFCHIHATCEYNSNGTASGICAKGYE
                     GDGTLCEMDPCTGLTPGGCSRNAECIKGTGTHTCVCQGWNGNGRDCSEINNCLLP
                     SAGCHDNASGLYVGPQNECECKKFRGNGIDCEPITSCLEQTKGCHPLASQSTSS
                     GVMSCVQEGYEGDGFCLCYNAAVELSLSEAAIFNRWINNASLQPLSATSNLTIV
                     PSQATBMDQDEKSWLSQSNIPALIKYHMLKTYEVADLQTLSSDMLATSLQGNF
                     LHAKVQNTITEGASVDGDNATNGVHIINKVLYPQRRLTSGSLFNLMRLEQMPD
                     YSIFRGYIIQYNLANAEADAYTVAFPNNAENIYREKVLVSLEEDLRYHVQLEE
                     KLLKNDLHNGHMETLGFYSFLFLHNDQLYVNEAPINYNTVATDKVHGLGKYL
                     EIQKRDNDNDTTIIRGCRITCSSELTCPFGTKSLGNEKRCRITVSYFMGRRTIFIG
                     QPKCVRTIVITRECCAGFPQCPGNAQVCFNGICLDGVNGTGVCEGEGFSGT
                     ACETCTEGYGIHCDQACSCVHGRCNQGPLDGSDCDVGWRGVHCDNATIEDNCNTG

```

## ORIGIN

## Alignment Scores:

Pred. No.: 3,46e-164 Length: 5604  
 Score: 1772.00 Matches: 334  
 Percent Similarity: 95.18% Conservative: 2  
 Best Local Similarity: 94.62% Mismatches: 17  
 Query Match: 95.01% Indels: 0  
 DB: Gaps: 0

US-09-466-778B-11 (1-353) x AK074051 (1-5604)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 Db 3984 ATGACAGCCCGGGCAGCACAGTGTGAGCTGTATAAGTCACTATGTCGAGAGTGGCTG 4043  
 Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 Db 4044 AACTGTGAGCGGAGCAGCTGCCCAITGACCGCTGTCTACAGGACAAATGGCAGTGCAT 4103  
 Qy 41 AlaAsnAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 4104 GCGACGCCAAATGTGTCGACCTCCATCCAGATACCACTGTGGGGGTGTCCATCTA 4163  
 Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 4164 CGCTCCCACTGGCGCAGTATAAGTGAACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 4223  
 Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
 Db 4224 GAAGCTGCGACCATGGCAACCTCAACACAGCTCTCTTATGCCAGAGAGGCCAAGTACCAC 4283  
 Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 4284 CTGTGCTCAGACGCTGGCTGGAGACCGGGCGGTGGCTACCCCAACAGAGCTTGGCTCC 4343  
 Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 4344 CAGAACTGTGGCTCTGTGTGTGGTATAGTGGACTATGGACCTAGACCTAGACCAACAGAGT 4403  
 Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
 Db 4404 GAAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCACCTGCAAGGTGGG 4463  
 Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180  
 Db 4464 TATGTGGAGATGGCTTCTCATCGATGGGAACCTGCTGCAGGTCTCTGATGCTCTCCC 4523  
 Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 Db 4524 TCACACAACTTCCTGACGGAGGTCTGGCTATTCCAAACAGCTCAGCTCAGGCGGT 4583  
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220



Db 4584 GCATTCTAGAACACCTGACTGCTCATCCCGGACCTCTTTGTGCACAGAAC 4643  
 QY 221 SerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisLeuAlaAen 240  
 Db 4644 AGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGAGCATCAGACACCTCGCAAT 4703  
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAenGlyThrThrLeuGlnThrArgLeuGly 260  
 Db 4704 GTCAGCATGTTTCTTACCAATGACCTTGTCAATGGCACCACCTGCAACAGGCTGGGA 4763  
 QY 261 SerLeuLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 Db 4764 ACCAGCTCTCATCATCTCCAGCCAGGACCACTCAACAGCAGGACCCAGGTTTGT 4823  
 QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 Db 4824 GATGAAGAGCCATCTTCAGTGGGACATCTTGGCTCCCAATGGGATCATTCATGCAIT 4883  
 QY 301 SerArg\*\*\*LeuLeuAlaProAlaProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 4884 TCCAGGCTTTAAAGACACCCCTGCCCCCGTGACCTTGACCCACACACTGGCTTGGAGCA 4943  
 QY 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340  
 Db 4944 GGGATCTCTTGGCCATCATCTCTGGTGACTGGGGCTGTGGCTTACTCTCTAC 5003  
 QY 341 PheArgIleAsnArgGlyThrIleGlyPhe\*\*\*HisPhe 353  
 Db 5004 TTTGGATAAACCGGAGAACATCGCTTCCAGCATTTT 5042

RESULT 4  
 AB052958  
 LOCUS AB052958 Homo sapiens mRNA for FELE-2, complete cds. linear PRI 11-SEP-2002  
 DEFINITION  
 ACCESSION AB052958  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Adachi, H. and Tsujimoto, M.  
 TITLE FEEL-1, a novel scavenger receptor with in vitro bacteria-binding and angiogenesis-modulating activities  
 JOURNAL J. Biol. Chem. 277 (37), 34264-34270 (2002)  
 MEDLINE 2206516  
 PUBMED 12077138  
 REFERENCE 2 (bases 1 to 8214)  
 AUTHORS Adachi, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-DEC-2000) Hideki Adachi, RIKEN, Cellular Biochemistry Laboratory, Hiroshima 2-1, Wako-shi, Saitama 351-0198, Japan  
 (E-mail: adachi@postman.riken.go.jp, Tel: 81-48-467-9372, Fax: 81-48-462-4670)

FEATURES  
 source  
 1..8214  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 gene 1..8214  
 /gene="FELE-2"  
 CDS 183..7838  
 /gene="FELE-2"  
 /note="fascilin, EGF and laminin type EGF containing scavenger receptor-2"  
 /codon\_start=1  
 /product="FELE-2"  
 /protein\_id="BAC15608.1"  
 /db\_xref="GI:22779441"  
 /translation="MMQLVIVFCLGLVQNFCSPAFTTGGARRCDKSLTLTIRCR SCALNLGVKPDGYMTITSSGVGRDCEYTFEVRTYLSLPGCRHICRDKYLQPRCP

## ORIGIN

## Alignment Scores:

Pred. No.: 5,6e-164 Length: 8214  
 Score: 1772.00 Matches: 334  
 Percent Similarity: 95.18% Conservatives: 2  
 Best Local Similarity: 94.62% Mismatches: 17  
 Query Match: 95.01% Indels: 0  
 DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x AB052958 (1-8214)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGluVleu 20  
 Db 6509 ATGACAGCCCGGCAAGCACAAGTGTGAGTGTAAAGTCACTATGTGCGAGATGGGCTG 6568  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAenGlyGlnCysHis 40  
 Db 6669 AACTGTGAGCGGAGCAGCTGCCCAITTGACCGCTGTACAGCAATGGCAGTGCAT 6728  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 Db 6729 GCAGAGCCCAATGTGCGACCTCCACTTCCAGATACCACTGTGGGGTGTTCATCTA 6788  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 Db 6789 CGCTCCCACTGGGCGAGTATAGCTGACCTTTGACAGCCAGAGGCGCTGTGCAAC 6848  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100

GRWGPDICPCGAGSPNGRSGCAEGMEGNTCSQEGGCTGACTACCTCADDNLFGPSC  
 SSVCNVGVNSGLDGTCTCYSAITGPKCDKPIPECAALLCPENRSPSPSTDEN  
 KLECKLNYRGDGYKDPINPLRKICHPHACTYLGPNRHSCTCQEGYRGDQVCL  
 PVDQINFGRCPTKYCKVDPQSCHECKEHTYQNFVPGVGGSMIDICKSDNPCR  
 NACTYVAGRTECICQGYVGDGLTCYGNIMERLELNRELNRELNRELNRELNRELN  
 KAYAWPLSKLGFVTLPTDKGLKFNVELVDNKAAYFVKLHI IAGQNI EYVNN  
 TDMFTYTLKSGEINFSNDQIKLKHGGKKVKIIQODI IASNGLHLILDRAMDKL  
 EHTFESNNEQITMTLQPRYSKFRSLLEETLHALDDEGCGPYTIFVPNNALNNM  
 KQTDYLLSPGSKLLELVYHIVPTOLEVATLI STPHRSNANQLIQNTTNDG  
 QILANDVAMELEITAKNGRYTLTGVLIPSVIFLPHRCDETRKMLGTCVCSCL  
 VYWRCPANPFTALTHRCSYVSGFSGKSGCARYCNATVKIPKCGYFGPDNCQC  
 PGFSPNSGNGQCADSLGNGTCTCEGFGSGQCFGSDPNKYGRNKKCLCVHGT  
 CNNRIDSGACTCTCRDGSAGRLCDKTSAGPYVQVCHIHATCEYNGTASCICCA  
 GYEGDGTLCSEMDPCTGLTGGCSRNAECIKTGTHTCVQOQGTNGRDCSEINNA  
 LLAGGCHGNASCLYVPGONECECKFRNGIDCEPITSCLEQTKCHPLASCQS  
 TSSGVMSVCCEGVEGDFLYGNAVELSFLSEAAIFNRNNASLOTULSATNL  
 VIVPSQATEMDQDEKSFWSQSNIPALIKIHMLLGTVRVADLOTLSSDMLATSLQ  
 GNFHLAKVDGNIIEGASIVEDGNAATNGVHIINKVLVFORLITGSLNLLMLEQ  
 MPDYSIRGYITIQYNLANAIEADATVFPANNNAIENYIREKKVLSLEEDVLRHV  
 LEEKLNDRHNGHRETMFGFSYPLSEFLNDOLYNEAPINTVNATDKGVHGLG  
 KVLLETQNRKDNNDTIIIRGRCTCSSLTCFPGTKSLGNEKRCIYTSYPMGRITLF  
 IGCOPKCVTVITRECCAGFPQCPQCPGNAQNVCFNGICLDGVNGTGVCECEGF  
 SGTACTCTEKYGIHCDQACSCVHRCNQGPDGSDCDVGRWGRHCDNATDNC  
 NGTCHTASNTSDGTASCKAAGFQNGTICTAINACEISNGCSACNCLPKRTTPG  
 RRVCTKAGYTGDIIVCLEINPLENHGGCDKNAECTQTPGNAACNCLPAYTGDKV  
 CTLINVCITKNGGSEFAICNHTQVETCTCKPNYIGDGTGCGSIVQELPKNPKTS  
 YFFOLOHFVKDLVGPFTVPAPLSADEFDEARVKDMDKVGMPQVLRHVYVACHQ  
 LLENKLISNATSLQGEPIVIVSQSYVINAKKILSSDIISTNGVHIHIDKLLSP  
 KLLITPKDSGRILQMLTLATNNGYIKFSNLIQDSGLLSVITDPIHTPTLFWPTD  
 QALHALPAEQDFFLNGQNDKDKLYLKFHVIRDAKVLADLPTSTAWKTQSGELV  
 KCGARDIGDLFNGQTCRIVORELLDGVAYIGDILLIDTLGRCDDTFTDASG  
 ECGSVANTPCPRMSKPKVKQKLYNLKPNKNGRCRCSLVIQIPRCCGYVGRD  
 COACGPGDAPNNRGVCLDQYSATGCKNGTNGTACEMWPCWPGDCLPGCSD  
 HGQCDGIGTSGOCLCETWGTGSCDQAVLPVCTPPCSHAHCKENNTCECNIDYE  
 GDGICHTVDFPKDNGGCAKVARCKGTGKVCSCQKGYGKDGHSCTEIDPCADGLN  
 GGCHETACKMTGPKHCKECKSHYVGDGLNCEPEQLPIDRCLQDQGCCHADKVDL  
 HFQDTVGVFLRSPLQYKLTFDKAREACANEATWATYNQLSYAQAKYHLCAGW  
 LETGRVAYTAFASQNGCGVGVIVDYPNPKSEMMWDFCYRMKDVNCTCKGVVGG  
 GFCSGNLLOVLMSFSLTNFLTVLAVSNSSAGRAFLBELHLDLSIRGLTFVQNSG  
 LGNETLSGRDIEHLNANSMFFYNDVNGTTLQRLGSKLLITASQDQPTSTFV  
 DGRALQWDIFASNGIHSRPLKAPPVTLTHLGAGIFFAILLVGTGAVALAY  
 SYFRINRTITGFQHFSEEDINVAALGKQKQPENISNPLYESITTSAPPEPSYDPTDSE  
 ERQLEGNDPLRLT"

Db 6849 GAAGTCGACCATGGCAACCTACCAACAGCTCTCTCTATGCCCGAAGGCCAAGTACAC 6908  
 Qy 101 LeuCySerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 Db 6909 CTGTCTCAGCAGGCTGGTGGAGACCGGGGGTGGCTTACCCACAGCTTCGCCCTC 6968  
 Qy 121 GlnAsnCyGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140  
 Db 6969 CAGAACTGTGGCTCTGTGTGGTGGATAGTGGACTATGGACCTTAGACCCCAACAGAGT 7028  
 Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCySerThr\*\*\*LysValGly 160  
 Db 7029 GAAATGTGGATGCTCTCTGCTATCGATGAAGATGTGAATGCACTGCAGGTGGCG 7088  
 Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGluValLeuMetSerPhePro 180  
 Db 7089 TATGTGGAGATGGCTTCTCATGCAGTGGGAACCTGCTCAGGTCTCATGTGCTTCC 7148  
 Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 Db 7149 TCACCTACAACTTCTGTACGGAAGTGTGGCTATTCACAGCTCAGCTCAGCGCCGT 7208  
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 Db 7209 GCATTTCTAGAACACCTGACTGACCTGCTCATCCGCGGACCTCTTTGTGCCACAGA 7268  
 Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 Db 7269 AGTGGCTGGGGGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCATCCCA 7328  
 Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGluThrArgLeuGly 260  
 Db 7329 GTACAGATGTTTTCTACAAATGACCTTGTCAATGGCACCACCTCGAAACGAGGCTGGA 7388  
 Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 Db 7389 AGCAAGCTGCTCATCTCCAGCCAGGACCCACCTCCACACCGAGAGACCATGTTGTT 7448  
 Qy 281 AspGlyArgAspThrLeuGluThrIlePheCysAlaSerAsnGlyIleThrHisValle 300  
 Db 7449 GATGAAGAGCCATCTCGAGTGGGACATCTTGTCTCCATGGGATCATTCATGCTCAT 7508  
 Qy 301 SerArg\*\*\*LeuLysAlaProAlaProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 Db 7509 TCCAGGCTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACACTGGCTTGGAGCA 7568  
 Qy 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340  
 Db 7569 GGGATCTTCTTGGCCATCATCTGCTGAGTGGGGCTGTGGCTTGGCTTACTTCTTAC 7628  
 Qy 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 Db 7629 TTTCGGATAAACCGGAGACAATCGGCTTCCAGCATTT 7667  
 RESULT 5  
 AY311388  
 LOCUS AY311388  
 DEFINITION Homo sapiens FEX2 mRNA, complete cds.  
 ACCESSION AY311388  
 VERSION AY311388.1 GI:32351284  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 8251)  
 Park,S.-Y. and Kim,I.-S.  
 FEX2, a novel cell adhesion molecule of Fas-1 superfamily mediates  
 cell-cell interaction  
 Unpublished  
 2 (bases 1 to 8251)  
 Park,S.-Y. and Kim,I.-S.  
 Direct Submission

JOURNAL Submitted (31-MAY-2003) Department of Biochemistry, School of Medicine, Kyungpook National University, 101 Dongin-dong, Jung-gu, Daegu 700-422, Korea  
 FEATURES  
 source  
 1. 8251  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /map="12q23.3"  
 205..7860  
 /note="FAS1, BGF-like and X-link domain containing adhesion molecule-2"  
 /codon\_start=1  
 /product="FEX2"  
 /protein\_id="AAP74958.1"  
 /db\_xref="GI:32351285"  
 /translation="MMLQHLVIFCLGLVQVFCSPAETTGQARRCDKSLILTIETCR  
 SCALNLGKCPDGYTMTSSGVGVRDCRYTEVRYTSLSLPGCHRICRKYLRCCP  
 GRWPDCTIECGGAGSPCNGRGSCAEGMEGNGTSCQEGFGGTACETCADDNLFGPSC  
 SSNCNVHGVSGLDGDTCECYSAITGPKDKPIPECAALLCPENSRSCPSFEDEN  
 KLECKLPNRYGDKYCDIPNPKRIKCHPAACTYLPNHRHSCCTCQGYRGDQVCL  
 PVDPCQNFNCPCTKTVKVDGPGQSCHECKHYQNFVPGVGMIDICKSDNPCHR  
 NACTTVAPGRTEGICOKYVYGDGLTCYGNIMERLRELTPRGKQGLTSLFSLLD  
 KAYAPLSKLPFTVLAPTDGLAGFNVELLDNKAQYFVKHLILAGOMNIEWMN  
 TDMYTTGSGSEIPNSDKONQIKLHGGKKVKITQGDIIASNLHLIDRAMDKL  
 EPTFESNNEQITMTLQPRYSKFRSLIEETNLHALDEDDGVGGPYTFVPPNEALNNM  
 KGDYLLSLSPGSKLLLELYRHYVFTFTEVATLSTPHIRMANQILQNTDNG  
 QILANDAMEIEITAKNGRIYTLTLPSPVILPHRDETKRKLGTGVCSSL  
 YVMRCANSEPTALTFRCHYVSGRFGSLKGCARYCNATVKIPKCKGFGVPCNOC  
 PGGSPNCNGGACDLSGGNGTICREFGOGSOGCPSPDNKYPNKKCLCVHGT  
 CNNRDSGACLTGTCRDGSGRGLCDKQTSAGGYVQFCHHATCEVNSGTASCIKA  
 GYEGDGLCSMDPCTGLTPGGCSRNAECIKGTGTCTVCVHCQQTNGRDCSEINN  
 TSSGAGCCNASCLYVPGQNECECKGRFNGIDCEPITSCEQTKCHPLASCS  
 LSSGVSCVCGEYGGDFLCYGAVALSLSBAAPNRWINNASLOPTLSAYSNLT  
 VLPSQATQEDMDDEKSWLSQSNIPALIKYHMLGTYRVADLQTLSSDMLATSLQ  
 QNFLHLAKVDNITIEGASIVDGNATNGVIHILNKILVLPQRLTSLNLLMRLEQ  
 MEDYSIFRGYIIQYNLANAIEADATVFAFNNAENIYREKVLSEEDVLRHVH  
 LEEKLNDLHNGMRETHRETFYSLFFLNDQLYNEAPINTVATDGVGHGLG  
 KVLBIQNRKNDNTTIRGCRITCSSELTCPFGTKSLGNEKRCIYTSYPMGRITLF  
 IGCPKCVRTVITRECCAGFPQCPQCPQNAQVFCNGICLQDVGNGTVCCEGEGF  
 SGTCETCEGKYHCDQACSCHVRCNQPLGDSGCDVGRGVRGCDNATTEDC  
 RVTCHTANCLTNSDGTASCKAFQNGTICTAINACEISNGCSAKADCKRTTPG  
 NRVCTKRGYTGDIYVCLINPCLEHGGCDKNAECTGTPNQACNCLPAYTGGKV  
 CTNLNCLTKNGGCSBFAICNHTQVETCTCKNYTGDFTKGSITQELPKPKTS  
 QYFQLOHFHVDLVLGGFPFTVPAPLSAFADEARVDMKDWKGLMPQVLRHVACHQ  
 LLLNKLINATSLQGEPIVISQSTVYINNKAIISSDIIISNGIVHIDLLSP  
 KNLITPKNSGRILQNLTLTANNYIKFSLNLIQDSGLLSVITDPIHTPLFWPTD  
 QALHALPABQDFLPNODNKLKYLKFKHVRDAKVLAVLPTSTAWKTQGSLSV  
 KCGAGRIDGLFLNGQCRIVORELLPDGLVAYGIDCLLIDPTLGGRCDTFTTDSG  
 ECGSVTPSPCPRMSKPKYQKCLYNLPKRNLEGCEERCSLVIQPRCKGYFGRD  
 CQACPGDPAPCNRRGCLDQYSAIGCKTNGFNGTACENMPCWGRGPDCLPCGSD  
 HGQCDGTITGGQCLCTGTMTGSPCQDTQAVLPVCTPPCSAHATCKENNTCECLDYE  
 GDGITVDFQDNGGCAKVARCSQKTVSCQKGYGDSHSTETIDPCADGN  
 GGCHHATCKMGPCKKCKSHVGDGLNCEPQPIDRCLQDGOCHADAKCVDL  
 HFQDTVGVHRLSPQYKLTFDKAEACAEATQALYNQLSVAOKAKVHLCSAGW  
 LETGRVAPYAFASQNGSGVGVIVDGPKNKSEMDVFCYRKNVDNCTCKGVYGD  
 GFSCGMLLQVLMFSPSLNFTLVAYSNSARGAPLEHLTDLISIRGLTUFQNSG  
 LGNETLSGRDIEHLANVSFFYNDLVNGTTLQRLGSKLLITASQDLPQPTFRFV  
 DGRAIQLWDIPASNGIITHVSRPLKAPPVLTHTGLGAGIFPAIILVTGAVAAAY  
 SYFRINRTTIGFHFSEEDINVAALGKQKQPSNPLTSTSTTAPPEPSYDPTDSE  
 ERQLEGNDPLRLT"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.63e-164 Length: 8251  
 Score: 1772.00 Matches: 334  
 Percent Similarity: 95.18% Conservative: 2  
 Best Local Similarity: 94.62% Mismatches: 17  
 Query Match: 95.01% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-466-778B-11 (1-353) x AY311388 (1-8251)

```
QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 6631 ATGACAGCGCGGCAAGCAAGAGTGTGAGTGTAAAGTCACTATGTGCGAGATGGCTG 6690
QY 21 AsnCysGluProGluGlnLeuProlLeuAspArgCysLeuGlnAspGlnGlyCysHis 40
Db 6691 AACTGTGAGCGGAGCAGCTGCCATTGACCGCTGTCTACAGGACCAATGGCGAGTGGCAT 6750
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 6751 GCGAGCGCCAAATGTGCGACCTCCACTCCAGGATACCACTGTGGGGTGTTCATCTA 6810
QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 6811 CGCTCCCACTGGCGCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 6870
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 6871 GAAGCTGCGACCATGGCAACCTCAACAGCTCTCTATGCCAGAGGCAAGTACCAC 6930
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 6931 CTGTGCTCAGAGGCTGGCTGGAGCGGGGGTGGCTTACCCACAGCCTTCGCTCC 6990
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
Db 6991 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGACTAGACCAACAGAGT 7050
QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 7051 GAAATGGGAGTGTCTCTCTATGCGATGAAAGATGTGAACCTGCAAGGTGGGC 7110
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyValAsnLeuGlnValLeuMetSerPhePro 180
Db 7111 TATGCGGAGATGGCTCTCATGTCAGTGGGAACTGCTGCGAGTCTGTATGCTCTCCC 7170
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerAlaArgGlyArg 200
Db 7171 TCACCTCACAACTTCCTGACGGAAGTCTGGCTTATTCACAGCTCAGCTCAGGCGCT 7230
QY 201 AlaPheLeuGluHisLeuThrAspLeuSerLeuArgGlyThrLeuPheValProGlnAsn 220
Db 7231 GCATTTCTGAACACCTGATGACCTCTGCTGGCGGACACCTCTTTTGGCACAGAAC 7290
QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240
Db 7291 AGTGGCTGGGAGAGATGAGACTTCTGCTGGCGGACATCGAGCACCACTCGCAAT 7350
QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 7351 GTCAGCATGTTTCTTACAAATGACCTTGTCAATGGCACCACTGCAACAGAGCTGGGA 7410
QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 7411 AGCAAGCTGCTCATCTGCGCAGCAGACCCCTCCACCGAGGAGACAGAGTTGTT 7470
QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 7471 GATGGAAGAGCCATTCTGCGAGTGGGACATCTTTGCCCTCCAATGGATCATTCATGTCAT 7530
QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 7531 TCCAGGCTCTTAAAGCACCCCTTCCCGCCGCTGACCTTGACCCACACCTGGCTGGAGCA 7590
QY 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
Db 7591 GGGATCTTCTTGCCATCACTCTGCTGAGTGGGCTGTGCTTGGCTGCTTACTCTCTAC 7650
QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7651 TTTCGGATAAACCGAGACCAATCGGCTTCCAGCATTTT 7689
```

```
RESULT 6
HSA295695
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
REMARK
FEATURES
source
```

```
HSA295695      9266 bp      mRNA      linear      PRI 28-JAN-2003
Homo sapiens mRNA for stabilin-2 (STAB2 gene).
AJ295695
AJ295695.1 GI:18073427
STAB2 gene; stabilin-2.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Politz, O., Gratchev, A., McCourt, P.A.G., Schledzewski, K.,
Guillot, P., Johansson, S., Birk, R., Hakiy, N., Franke, P., Kodelja, V.,
Kannicht, C., Orfanos, C.E., Johansson, S. and Goerd, S.,
Longati, P., Velden, F.W., Johansson, S. and Goerd, S.,
Stabilin-1 and -2 constitute a novel family of fasciclin-like
hyaluronan receptor homologues
Biochem. J. 362 (Pt 1), 155-164 (2002)
21818465
11829752
Politz, O., Gratchev, A., McCourt, P.A.G., Schledzewski, K.,
Guillot, P., Johansson, S., Birk, R., Hakiy, N., Franke, P., Kodelja, V.,
Kannicht, C., Orfanos, C.E., Johansson, S. and Goerd, S.,
Stabilin-1 and stabilin-2 constitute a novel family of fasciclin
domain-containing adhesion molecules associated with
endothelial-macrophage differentiation and angiogenic processes
Unpublished
3 (bases 1 to 8266)
Gratchev, A.
Direct Submission
Submitted (27-JUL-2000) Gratchev A., Department of Dermatology,
University Medical Center Mannheim, Theodor-Kutzer-Ufer 1, D-68135
Mannheim, GERMANY
Revised by author 27-APR-2001
Location/Qualifiers
1..8266
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..8266
/genes="STAB2"
187..7842
/genes="STAB2"
/codon_start=1
/product="stabilin-2"
/protein_id="CAC82105.1"
/db_xref="GI:18073428"
/db_xref="GOA:Q8WWQ8"
/db_xref="SPTREMBL:Q8WWQ8"
/translat="MMLOHLVIFCLGLVVFQNFCSPAETTGQARRCDRKSLILITRTCR
SCALNLGVKPDGYTMTSGSVGRDRTYTFVRYTSLSLGCRHICRKYDLQPRCP
GWGPCICPGAGSPGNGRSCGMEGNGTCSCQGGGTGCTACDNLQPRCP
SSVQNCVHGVCNSGLDGDCTCYSAYTGPCKDKPIPECAALICPENSRCSPSTEDN
KLECKLPNTRGDKYDIPNCLRKICHPHAKHTYLGPNRHSCPTQEGYRGQGVCL
PVDPQINFGNCTKSTVKYDGPQSHCEHYQNFVPGVSCMTDICKSNCPCH
NANTCTVAPRTETCICQKGVGDLTCYGNIMERLENTPEPRGQWGLTSLD
TMFYTLTKGSGEIPNSKDNOIKLHGKGVKVIQDIIASNGLLHIDRAMDKL
KAWPLSLGPFVLLPTDKGLGFNVNELLVDNKAQYFVKLHIAQMNIEYNN
EHTFESNQTINTMLQPRYSKFRSLLENTNLGHALDEGVGPPYTI FVNNFALNM
KUTLDYLLSPESRKLILVRYHVPTQLVATLISPHIRSMANQLIQNTVING
QILANDVAMEIEITAKNGRIYTLTLPPSIVPLPHRDETKREMLGTCVCSGL
VYWSRPNANSEPTALFTRCVYSGRGLSKGCARYCNATVKIPCKGFGYGDQCQ
PGFSNPGSCNGQCADLGGNGTCICEGFGSQSCQFCDPNPKYRCKKXKLCVHGT
CNRNIDCDACLTCTCRDGSAGRLCDKQTSACGPYVQFCHIHATCEYSGTASCIKA
GYEGDCTLCSEMDPCTGLTPGSCSRNAECIKTGTHGTCVCOQWNGNRDCSEIINC
LLPSAGCGCHDNASCLYVPGQNECECKGFRNGIDCEPITSCLEGTGKCHPLASQCS
LVSPGWSVCVCGEGDGLCYGNAAVELSFSEAAIFNRINNALSQTLTSATSNLT
VLVPSQATDEMDQDEKFWLSQSNIPALIKYHMLLGYRVADLTSLSSDMLATSLQ
GNPLHAKYDGIITIEGASIVDGNAAATNGVTHINKVLVQPRRLTGLSPNLRLMEQ
MPDYPIRGVYIIOYNLANAIEADAVTFAPNNAENIYREKKVLSEEDLVRYHV
LEEKLLKNDLHNGHRETMLGFSYFLSFFLHNDQLYNEAPINYTNVATDKVHGILG
KVLIEQKNRCDNNAITLIRGCRKTSSEUTCPFGTKSLGNEKRCRIYTSYFMGRRLIF
```





```
Db 1875 GAAGCTGGACCATGGCAACCTACCAACAGCTCTCTATGCCCCGAAGCCCAAGTACCAAC 1934
QY 101 LeuCySerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 1935 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGGGTGGCTTACCCACAGCCCTTGCCTCC 1994
QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 1995 CAGAACTCGGCTCTGTGTGGTGGTAGTGGAGTATGACCTATGACCTAGACCCCAACAAGAGT 2054
QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 2055 GAAATGTGGATGCTCTCTGCTATCGGATGAAGATGGAATGCACCTGCAAGGTGGGC 2114
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 2115 TATGTGGAGATGGCTTCTCATGAGTGGAACTGTGCTGCAGGTCTGTATGCTTCCCC 2174
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 2175 TCACCTACAAAATTCTCTCAGCGAAGTGTGGCTATTCCAAACAGCTCAGCTCGAGGCCGT 2234
QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 2235 GCATTTCTAGAACACCTGACTGACCTGTCCATCCCGCGGACCCCTCTTTGTGCCACAGAAC 2294
QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
Db 2295 AGTGGCTGGGGAGAAATGAGACCTGTGTGGCGGGACATCGAGCACCACTCGCCCAAT 2354
QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGluThrArgLeuGly 260
Db 2355 GTCAGCATGTTTCTACAAATGACCTGTGCAATGGCACCACTCCCAACAGCGGGGTGGGA 2414
QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 2415 AGCAAGCTGCTCATCCTGCGAGCCAGGACCCACCTCCCAACGACGAGACCAAGGTTGTT 2474
QY 281 AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 2475 GATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCTTCCATGGGATCATTCATGTCATT 2534
QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 2535 TCCAGGCTTTTAAAGACACCCCTCCCTCCCGCTGACCTTTGACCCACACTGGCTTGGAGCA 2594
QY 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
Db 2595 GGGATCTTCTTTGGCATCATCTGTGTGACTGGGGCTGTGCTTGGCTGCTTACTTCCCTAC 2654
QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 2655 TTTGGATAAACCGGAGAACAAATCGGCTTCCAGCATTTT 2693

RESULT 9
AX149461
LOCUS AX149461 3625 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 17 from Patent WO0136638.
ACCESSION AX149461
VERSION AX149461.1 GI:14347936
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Shimkets,R.A., Lichenstein,H., Vernet,C. and Fernandes,E.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0136638-A 17 25-MAY-2001;
Curagen Corporation (US)
FEATURES
source 1..3625
```

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 2.48e-143 Length: 3625
Pred. NO.: 1558.00 Matches: 300
Score: 85.31% Conservative: 2
Percent Similarity: 84.75% Mismatches: 15
Best Local Similarity: 83.54% Indels: 37
Query Match: 6 Gaps: 1
DB: 1

US-09-466-778b-11 (1-353) x AX149461 (1-3625)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 2078 ATGACAGGCCCGGGCAAGCACAGTGTGAGTGTAAAGTCACTATGTCCGAGATGGGCTG 2137
QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 2138 AACGTGTAGCGCGGAGCAGCTGCCCATTTGACCGCTGCTTACAGGACAAATGGGCGAGTGCAT 2197
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 2198 GCAGACGCCAAAATGTGTGCACCTCCACTTCAGGATACCACTGTGTGGGTGTTCCTACTA 2257
QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 2258 CGTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGGCTGTGCCAAC 2317
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 2318 GAAGCTCGACCATGGCAACCTACACAGCTCTCTATGCCAGAAAGGCCAAGTACCAAC 2377
QY 101 LeuCySerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 2378 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGGGTGCTTACCCACAGCCCTTGCCTCC 2437
QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 2438 CAGAACTGTGGCTCTGGTGTGGTATAGTGGACTATGACCTAGACCCCAACAAGAGT 2497
QY 141 GluMetThrAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValGly 160
Db 2498 GAAATGTGGATGTCTTCTGCTATCGGATGAAAG- 2531
QY 160 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePr 180
Db 2531 ----- 2531
QY 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
Db 2532 -----GAAGTGTGGCTATTTCACACAGCTCAGCTCGAGCCG 2569
QY 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
Db 2570 TGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCAACCCTCTTTGTGCCACAGAA 2629
QY 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAs 240
Db 2630 CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCACTTGGCAA 2699
QY 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGl 260
Db 2690 TGTGAGCATGTTTTCTTACAATGACCTTGTCAATGGCACCACTCCCTGCAACAGAGCTGGG 2749
QY 260 ySerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 2750 AAGCAAGCTGCTCATCTACCTCCAGCCAGGACCCCACTCCAAACCGGAGGACCAAGTTGT 2809
QY 280 lAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIl 300
Db -----
```

Db 2810 TGATGGAAGCCATTCTCGAGTGGGACATCTTTGGCTCCAAATGGGATCATCTCATGTGCAT 2869

Qy 300 eSerArg\*\*\*LeuylsAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\* 320

Db 2870 TTCCAGGCGCTTTAAAGACACCCCTCGCCCGCTGACCTTGACCCACACTGGCTGGGAGC 2929

Qy 320 \*GlyLeuPhe\*\*\*\*\*llelleuValThrGlyAlaValAlaLeuAlaValThrSerTy 340

Db 2930 AGGATCTTTCTTGGCCATCATCTCGGTGACTGGGGCTGTGGCTGGCTGCTTACTCCTA 2989

Qy 340 rPheArgIleAAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353

Db 2990 CTTTCGGATTAACCGGAGACATCTGGCTCCAGCATTTT 3029

RESULT 10

HSMB01377

LOCUS

DEFINITION Homo sapiens mRNA; cDNA DKFZp434E0321 (from clone DKFZp434E0321).

ACCESSION AL133021

VERSION AL133021.1 GI:6453577

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS Blum, H., Bauersachs, S., Mewes, H. W., Gassenhuber, J. and Wiemann, S.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-1999) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by IMU (Ludwig Maximilians University, Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp434E0321) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source

1. .3642

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp434E0321"

/tissue\_type="testis"

/clone\_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"

/dev\_stage="adult"

join(7..2607,2609..3217)

/gene="DKFZp434E0321"

join(<7..2607,2609..3217)

/gene="DKFZp434E0321"

/note="strong similarity to KIAA0246

The frame shift was determined manually"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="CAB61358.1"

/db\_xref="GI:6453578"

/db\_xref="GOA:Q9UF98"

/db\_xref="SPTRMBL:Q9UF98"

/translation="VGRVAVTASCKCAAGFGNGTICTAINACEISNGGCSAKADCKR

TTGRRVCTKAGTVDGIVLCINPCLENHGGCDKNAECTQGPNOACNCLPVTG

DGKVTILNVLKNGCSFAICNHTGOVETCTCKPNYIGDFTGRGSIYDELPHN

PKTSQYFQIQEHFVLDVGPFTVAPLSAADFBEARVKDWKGLMPQVLRKYHV

ACHOLLLENKLLSNATSLQGEPIVISVOSTVYINNKAIISSDIISTGIVHIDK

LLSPKNIILTPKNSGRILNLTTLATNNGYIKFNSLLQDSGLLSVITDPIHPVILF

WPTDQAULHAPASQDPLFNQDNKDKLKEYLKHVIRDAKVLAVDPTSTAWKTLQGS

ELSVKCGAGRIDGLFLNGQTCIRVQRELLFDLGVAYGIDCLLIDPTLGRCDTFTF

DASGECGSVNTDSCRPWSKPGVKQKCLNLPKRLNLEGRCRCSLVIQIPRCCKGY

FGRDCAQCPGPPAPCNRRGVCLDOYSATGECKNTGNCGTACEMCWPFGRFGDCLPC

GCSDHGQDDGIGTSGGQCLTETGWTGSPSCDTQAVLPVCTPPCSAHATCKENNTCEGN

LYVEGDICTTVDFCKQNGGCAKVARCSQKTKVSCSQKVKYKGDHSCTEIDPCA

DGLNGCHEHATCKMTGPKHCKECSHYVGLNCEPEQLPIRDLQNGQCHADAK

CYDLHFQDITTVGVFHLRSPLQGYKLTFDKAREACANEATMATYNQLSYAKAKYHLC

SAGLWETGVAVPTAFASQNGSVVGIVDYGRPNKSEMDVFCYRMKSGSAGLFOOL

SRPCTSRTPDDLSIRGTLFVQNSGLGENETLSGRDIEHHLIANSMFVYNDLVNGET

LQTRLSKLLITASDPLQPTETRFVDSGRALLQWDIPASNGLIHVSRLKAPPAVPT

LHTHGGAGIFFAILLVTAVALAAYSIFRINRTIGFQHFSESEEDINVAALQKQOPE

NISNPLEYESTTSPPEPSYDPFTDSEERQLEGNDRPLRLT"

3574..3579

polyA signal

polyA\_site

ORIGIN

Alignment Scores:

Pred. No.: 2,49e-143 Length: 3642

Score: 1558.00 Matches: 300

Percent Similarity: 85.31% Conservative: 2

Best Local Similarity: 84.75% Mismatches: 15

Query Match: 83.54% Indels: 37

DB: 9 Gaps: 1

US-09-466-778B-11 (1-353) x HSM801377 (1-3642)

Qy 1 MetThrGlyProGlyVlyeHisCysGluCysAlaSerHisTyrValGlyAspGlyLeu 20

Db 2095 ATGACAGGCCCGGCGAAGCAAGTGTGTGATGTAAAGTCACTATGTGCGAGATGGCTG 2154

Qy 21 AsnCysGluProGluInLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40

Db 2155 AACTGTGAGCGGAGCAGCTGCCCAITGACCGCTGTACAGCAATGGCAGTGCAT 2214

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60

Db 2215 GCAGAGCCCAATGTGTGACCTCCACTTCCAGGATACCACTGTGGGGTGTTCATCTA 2274

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80

Db 2275 CGTCCCGACTGGCGCAGTATAAGCTTGACCTTTGACAAAGCCAGAGGCGCTGTGCCAAC 2334

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100

Db 2335 GAAGCTGCGACCATGCAACCTACACAGCTCTCTCTATGCCCGAGAGGCCAAGTACCAC 2394

Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaLysThrProThrAlaPheAlaSer 120

Db 2395 CTGTGCTCAGCAGCTGGCTGGAGACGGCGGGTGGCTTACCCACAGCTTTCGCTCC 2454

Qy 121 GlnAsnCysGlySerGlyValValValGlyVileValAspTyrGlyProArgProAsnLysSer 140

Db 2455 CAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGACTATGGACCTAGACCCACACAGAGT 2514

Qy 141 GluMetTrpAspValPheCysTyrArgMet-LysAspValAsnCysThr\*\*\*LysValGln 160

Db 2515 GAAATGTGGGATGCTCTTCTGCTATCGGATGAAAG----- 2548

Qy 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr 180

Db 2548 ----- 2548

Qy 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200

Db 2549 -----GAAGTGTGGCTATTTCACAGCTCAGCTCAGTCGAGGCGG 2586

Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220

Db 2587 TGCATTTCAGAACACCTGACTGACCTCCATCCCGCGGACCCCTCTTTGTGCGCACAGAA 2646

Qy 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240

Db 2647 CACTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGAGCATCGACACACACCTGCCCA 2706

Qy 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGln 260

Db 2707 TGTACAGCATGTTTTCTACAAATGACCTTGTCAATGACCAACCCTCTCAACAGAGGCTGGG 2766



```

Qy 260 ySerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 2767 AAGCAAGCTGCTCATCTGCGAGCAGGACCACTCCAAACCGAGGACCAAGGTTTGT 2826
Qy 280 lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisVal11 300
Db 2827 TGATGGAGAGCCATTGCGAGTGGGACATCTTTGCCCTCCCAATGGGATCATTCATGTCAT 2886
Qy 300 eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
Db 2887 TTCCAGGCGCTTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACACTGGCTTGGGAGC 2946
Qy 320 *GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIleThrSerTy 340
Db 2947 AGGGATCTTCTTTGCGATCATCTCTGGTACCTGGGCTGTTGCCCTGGCTGCTTACTCTTA 3006
Qy 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 3007 CTTTCGGATAAACCGAGAACAAATCGGCTTCAGCATTTT 3046

RESULT 11
AX686194 8495 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 3 from Patent WO02059315.
DEFINITION AX686194
ACCESSION AX686194
VERSION AX686194.1 GI:29372027
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
Baldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,
Stone,D., Sciore,P., Millet,I. and Rothenberg,M.
TITLE Human nucleic acids and polypeptides and methods of use thereof
JOURNAL Patent: WO 02059315-A 3 01-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1..8495
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7,27e-143 Length: 8495
Score: 1558.00 Matches: 300
Percent Similarity: 85.31% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 15
Query Match: 83.54% Indels: 37
DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x AX686194 (1-8495)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyValGlyAspGlyLeu 20
Db 6948 ATGACAGGCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGGCTG 7007
Qy 21 AnCysGluProGluGluLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 7008 AACTGTGAGCGGAGCAGCTGCCCATGACCGCTGCTTACAGGACAATGGGCGAGTGCAT 7067
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 7068 GCAGACGCCAATATGTGCGACCTCCACTTCCAGGATACCACTGTGGGGTGTCCATCTA 7127
Qy 61 ArgSerProLeuGlyGlnTyrlLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 7128 CGCTCCCCACTGGGCCAGTATAAGTGTGACCTTTGACMAAGCCAGAGAGCGCTGTGCCAAC 7187

```

```

Qy 81 GluAlaAlaThrMetAlaThrTyAsnGlnLeuSerTy***GlnLysAlaLysTyHis 100
Db 7188 GAAGCTCGACCATGGCAACCTACCAACAGCTCTCTTATCCCAAGAGGCGCAAGTACCA 7247
Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaIleThrProThrAlaPheAlaSer 120
Db 7248 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGTGGCTTACCCCAAGAGCTTGGCTTCC 7307
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 7308 CAGAACTGTGGCTGTGGTGTGGATAGTGGACTATGGACCTACGACCAACCAAGAGT 7367
Qy 141 GluMetTrpAspValPheCysTyArgMet-LysAspValAsnCysThr***LysValGI 160
Db 7368 GAAATGGGATGTCTTCTGCTATCGGATGAAAG----- 7401
Qy 160 yTyrValGlyAspGlyPheSerTySerGlyAsnLeuGlnValLeuMetSerPhePr 180
Db 7401 ----- 7401
Qy 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaIleTySerAsnSerSerAlaArgGlyAr 200
Db 7402 -----GAAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGCGG 7439
Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleAtrGlyThrLeuPheValProGlnAs 220
Db 7440 TGCATTTCTAGAACACCTGACTGACCTGCTCCATCCCGCGGACCCCTTTGTGCCACAGAA 7499
Qy 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyValArgAspIleGluHisLeuAlaAs 240
Db 7500 CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGAGATCGAGCACACCTCGCCAA 7559
Qy 240 nValSerMetPhePheTyAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGI 260
Db 7560 TGTGAGCATGTTTTTCTACAAATGACCTTGTCAATGGCACCACTCGCAACAGAGCTGGG 7619
Qy 260 ySerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 7620 AAGCAAGCTGCTCATCTGCGAGGACCCCTCAACCGGAGGACCAAGGCTTGT 7679
Qy 280 lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisVal11 300
Db 7680 TGATGGAGAGCCATTCTGCGAGTGGGACATCTTTGCTCCCAATGGGATCATTCATGTCAT 7739
Qy 300 eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
Db 7740 TTCCAGGCGCTTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTTGGGAGC 7799
Qy 320 *GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIleThrSerTy 340
Db 7800 AGGGATCTTCTTTCGCCATCATCTGCGAGTGGGCTGTTGGCTTGGCTGCTTACTCTTA 7859
Qy 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7860 CTTTCGGATAAACCGAGAACAAATCGGCTTCAGCATTTT 7899

RESULT 12
AX686192 8444 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 1 from Patent WO02059315.
DEFINITION AX686192
ACCESSION AX686192
VERSION AX686192.1 GI:29372026
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
Baldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,

```



Stone, D., Sciore, P., Millet, I. and Rothenberg, M.  
Human nucleic acids and polypeptides and methods of use thereof  
Patent: WO 02059315-A 1 01-AUG-2002;  
Curagen Corporation (US)

# FEATURES

Location/Qualifiers  
1..8444  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
Pred. No.: 1 79e-142 Length: 8444  
Score: 1554.00 Matches: 296  
Percent Similarity: 85.27% Conservatives: 5  
Best Local Similarity: 83.85% Mismatches: 16  
Query Match: 83.32% Indels: 36  
DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x AX586192 (1-8444)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
Db 6907 ATGACAGCGCCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 6966  
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAsnGlyGlnCysHis 40  
Db 6967 AACTGTGAGCGGAGCAGCTGCCATTCACCGCTGCTTACAGGACAATGGGCGAGTGCCT 7026  
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
Db 7027 GCACAGCGCAATGTGCGAGCTCCACTTCCAGGATACCTGTGGGGTGTTCATCTA 7086  
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
Db 7087 CGCTCCCACTGGGCGCATTAAGTGTACCTTTGACAAAGCCAGAGAGCTGTGCCAAC 7146  
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*\*GlnLysAlaLysTyrHis 100  
Db 7147 GAAGCTGGGACCATGGCAACCTACACCAAGCTCTCTATGCCAGAGAGCCAAAGTACCAC 7206  
Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
Db 7207 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGTGTCTACCCACAGCTTGCCTCC 7266  
Qy 121 GlnAsnCysGlySerGlyValValGlyLeuValAspTyrGlyProArgProAsnLysSer 140  
Db 7267 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGACCTATGGACCTAGACCCCAACAGAGT 7326  
Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*\*LysValGly 160  
Db 7327 GAAATGTGGATGTCTTCTGCTATCGGATGAAG----- 7359  
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180  
Db 7359 ----- 7359  
Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
Db 7360 -----GAAGTGTGGCTATTCCACAGCTATCCAGCTCAGCTCGAGCGGT 7398  
Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
Db 7399 GCATTTCTAGAACACCTGACTGACCTGCTCCATCCCGCGGACCCCTCTTTGTGCCACAGAAC 7458  
Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAsnIleGluHisLeuAlaAsn 240  
Db 7459 AGTGGCTGGGGGGAATGAGACCTTGTCTGGCGGGACATCGACACACCTCCCAAT 7518  
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
Db 7519 GTCCAGCATGTTTCTACATGACCTGTGTCATGGCACCACCTCCGAAACGAGGTGGGA 7578

Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
Db 7579 AGCAAGCTGTCTATCACTGCACGACGAGCCCACTCCAAACGGGTACAAAGTAGTGTGTT 7638  
Qy 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
Db 7639 GATGGAGAGCCATTCTGCAGTGGGACATCTTGGCTCCATGGGATCATCATGTCAAT 7698  
Qy 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
Db 7699 TCCAGGCGCTTTAAAGCAGCCCTGCCCCGCTGACCTTGACCCACACCTGGCTGGGAGCA 7758  
Qy 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340  
Db 7759 GGGATCTCTTTCATCATCTCTGGTGTGAGTGGGGCTGTGGCTTGGCTTACTCTCTAC 7818  
Qy 341 PheArgIleAsnArgTyrThrIleGlyPhe\*\*\*HisPhe 353  
Db 7819 TTTCCGATAAACCGGAGACAATCGGCTACGACATTTT 7857  
RESULT 13  
AY007370 4706 bp mRNA linear ROD 23-OCT-2002  
LOCUS Rattus norvegicus hyaluronan receptor for endocytosis HARE  
DEFINITION precursor, mRNA, partial cds.  
ACCESSION AY007370  
VERSION AY007370.1 GI:24285892  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 4706)  
AUTHORS Zhou, B., Weigel, J. A., Saxena, A. and Weigel, P. H.  
TITLE Molecular Cloning and Functional Expression of the Rat 175-kDa  
Hyaluronan Receptor for Endocytosis  
JOURNAL Mol. Biol. Cell 13 (8), 2853-2868 (2002)  
MEDLINE 22169209  
PUBMED 12181351  
REFERENCE 2 (bases 1 to 4706)  
AUTHORS Zhou, B., Weigel, J. A. and Weigel, P. H.  
TITLE Direct Submision  
JOURNAL Submitted (25-AUG-2000) Department of Biochemistry and Molecular  
Biology, University of Oklahoma Health Sciences Center, 940 Stanton  
L. Young, BMSB 860, Oklahoma City, OK 73104, USA  
FEATURES  
1..4706  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
<1..4296  
/codon\_start=1  
/product="hyaluronan receptor for endocytosis HARE  
precursor"  
/protein\_id="AAG13634.1"  
/translation="SLESLTRLEQMPDYSIFRGYIHYNLAIESADAYTVFVFN  
EAIENYREKATSLKEDILRYHVVLGKLLKNDLHNGMRETMFGSYLLAFERND  
OLYNEAPINVTNATDKVHGLEKLEIKRDNNDTIIVRGEGKCSQAPCP  
ETKPLRETRKCIYSIYFMKRSVFIGQPCQVRIITRACCAAGFPQCAQCGRGN  
VCSNGCDVGVNGTGTCCGLGNGTACTCTEGKYGICHDQACSCVHRCSCQPLG  
AINACTSNCGSCPKADKRTTGNRVCKAGTGDGIVCLEINPCLNEHGGCDRA  
ECTGTGNQAVCNCLPKYTGKVCSLINCLTNNGCSPFAFCNTEQOQRICTCRP  
DYTDGIVCRGSIYVGLPKNPSTSYFFOLQEHAVRELAPGPTVPAFLSSFNHP  
RIKWDQGLQMSQVLRVHVVCQQLLDLLITPFDALGRVLQNLTVAAHNGYTKSLI  
AKVLSGDIISTNGVHVIDKLSPLNLLITPKDALGRVLQNLTVAAHNGYTKSLI  
QSGLSLVIIDSITHPTVFWPTDKALEALPPEQDFLNQDNKDKLAKYKLFVIR  
SKALASLPRASWKTQSGELSVRCGTGSDIGELFNEQCRFIHRLGLFDVGVNG  
IDCLMNPITLGGRCDTFTTIDIPOECGSCIPTPKPLKSPKPKKKCIYNPLPFRN  
VEGQCLNCTVIQTPRCHGYFMPDCCACPGPBDTPCNRNRCRDLVTPMQCLCHTG

FNGTACELCHWGRFGPDGCPDRSCSBHQCDEGITSGBCLCETGWTAAACDTPPTAVFA  
 VCTPACSVHATCENNTVCNLYEDGTCITVDFCKONNGGCKAKVAKSKOGKQTVS  
 CSCKGYKGDGYSCEIIEPCADGVNGGCHETACRTMGKHKCECKSHYVDGVDCE  
 PEQLPLDCLQNGQCHPDASCADLYFODTTVGVFHLRSPGLQYKLTDFDKAKEACAKE  
 AATATYCNOLSYAKAKYHLCSGRLVAYPTTYASOKCANVGVIVDYSRANK  
 SEMWDVFCYRMKDVNCTKAGYVGDGFCSGNLOVLMSFPSTLNFLEVLAFSKSA  
 RGOAFKLHLDLSIGTLVPONSGLPNGKLSGRDIEHLLTNVNVSFNDLVNGTFL  
 RMLSGQLLITVSQDLQHETRFVDRSILQMDIIAANGILHIIIEPUPAPPTAATAA  
 HSLGSGTFCVAVLVGTALAAAYSFLRKQRTTGFDFDQKTLMSMLLASSPRIS  
 QTLQWRPQRHPQSPVPVPSQTLNRIWRTATLWGHCGDMSRQQATTVTVR"  
 1..4293  
 mat\_peptide

/product="hyaluronan receptor for endocytosis HARE"  
 /note="175 kDa; derived by proteolytic cleavage"

## ORIGIN

## Alignment Scores:

Pred. No.: 6,69e-139 Length: 4706  
 Score: 1514.50 Matches: 280  
 Percent Similarity: 87.54% Conservative: 29  
 Best Local Similarity: 79.32% Mismatches: 43  
 Query Match: 81.21% Indels: 1  
 DB: 10 Gaps: 1

US-09-466-778B-11 (1-353) x AY007370 (1-4706)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 DB |||||  
 QY 3022 ATGACGGGCCCCAGGCAAGCATAGTGTGAATGTAAAGTCACATATGTCGGGACCGAGG 3081  
 DB |||||  
 QY 21 AenCysGluProGluGlnLeuProLysAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 DB |||||  
 QY 3082 GACTGTGACCTGAGCAGCTGCGGCTGACCGTGTCTACAGACACACGACAGTGCAC 3141  
 DB |||||  
 QY 41 AlaAspAlaLysCysValAspLysHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 DB |||||  
 QY 3142 CCAGATGCGCAGCTGTGCAGACCTCTACTTCCAGGACACGACCGTAGGATATCCACTA 3201  
 DB |||||  
 QY 61 ArgSerProLeuGlyGlnTrpLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80  
 DB |||||  
 QY 3202 CCGTCCCACTGGCGGAGTACAACTGACATTTGACAAAGCCAAAGACCTGTGCCAAA 3261  
 DB |||||  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*GlnLysAlaLysTyrHis 100  
 DB |||||  
 QY 3262 GAAGTGTGGTGCACCATAGCCACTACACAGCTCTCTATGCCAGACGACCAAGTATCAC 3321  
 DB |||||  
 QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 DB |||||  
 QY 3322 CTGTGCTGCGCGGCTGGTGGAGAGTGGCGGGTGGCTTACCCGACTACGTATGCTCTCT 3381  
 DB |||||  
 QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140  
 DB |||||  
 QY 3382 CAGAAGTGTGGTGCACCATAGCTGTGGATCGTAGACTAGGATCCAGGCCAACAGAGT 3441  
 DB |||||  
 QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*LysValGly 160  
 DB |||||  
 QY 3442 GAAATGTGGATGTCTCTGTACCCGATGAAAGATGTGAATGACCTGCAAGCAGGCG 3501  
 DB |||||  
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180  
 DB |||||  
 QY 3502 TATGTGGAGATGGCTTCTGTGGAGTGGGAACCTGCTGCAGGTCTCATGTCTCC 3561  
 DB |||||  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerAlaArgGlyArg 200  
 DB |||||  
 QY 3562 TCGCTCACAACTTCTGTACAGAGTGTGGCTTTTCCAAAGAGTCAAGGCGGAGAGCAG 3621  
 DB |||||  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 DB |||||  
 QY 3622 GCATTTTGAACACCTGACTGACTGCTCATCTCGTGGCAGCCCTGTTGTGCCACAGAAC 3681  
 DB |||||  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240  
 DB |||||  
 QY 3682 AGTGGGCTACCGGGAAATAAGAGCGTGTCTGCCCGGAGCATTTGAGCACCACTCACTAAT 3741  
 DB |||||

QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 DB |||||  
 QY 3742 CTCACAGCTCTCTTTTACATGACCTTCTCAATGTGTACCTTCTGAGGACTATGCTGGGA 3801  
 DB |||||  
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 DB |||||  
 QY 3802 AGCCAACTGCTCATTTACCTTCAGCCAGGACGAGCTCCAC--CAAGAGACGAGGTTGTG 3858  
 DB |||||  
 QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300  
 DB |||||  
 QY 3859 GATGGAAGATCATCTTCAGTGGGACATCATCGCGCAATGGAATCCTCCATATTATT 3918  
 DB |||||  
 QY 301 SerArg\*\*\*LeuLysAlaProProAlaProValThrLeu\*\*\*HisThrGlyLeuGly\*\*\* 320  
 DB |||||  
 QY 3919 TCTGAACCTTTGAGAGCTCTCCACGCGCAGCAACGGCTGCCCACTCTGGCGCTGGGACA 3978  
 DB |||||  
 QY 321 GlyIlePhe\*\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340  
 DB |||||  
 QY 3979 GGTATATTCTGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4038  
 DB |||||  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 DB |||||  
 QY 4039 TTCGGCTAAAGCAGCAGCAACCACTGTTTCCAGCGTTT 4077  
 DB |||||

RESULT 14  
 AF364951 8157 bp mRNA linear ROD 25-MAR-2002  
 LOCUS Mus musculus stabilin-2 mRNA, complete cds.  
 DEFINITION AF364951  
 ACCESSION AF364951  
 VERSION AF364951.2 GI:19705588  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 8157)  
 AUTHORS Polit,O., Gratchev,A., McCourt,P.A.G., Schledzewski,K.,  
 Guillot,P., Johansson,S., Svineng,G., Franke,P., Kamnitsch,C.,  
 Kzyshkowska,J., Longati,P., Veltan,F.W., Johansson,S. and  
 Goerdts,S.  
 TITLE Stabilin-1 and -2 constitute a novel family of fasciclin-like  
 hyaluronan receptor homologues  
 JOURNAL Biochem. J. 362 (Pt 1), 155-164 (2002)  
 MEDLINE 21818465  
 PUBMED 11829752  
 REFERENCE 2 (bases 1 to 8157)  
 AUTHORS Guillot,P., Schledzewski,K., Gratchev,A. and Goerdts,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAR-2001) Dermatology, Universitätsklinikum Mannheim,  
 Universitat Heidelberg, Theodor Kutzer Ufer 1-3, Mannheim, BW  
 68135, Germany  
 REFERENCE 3 (bases 1 to 8157)  
 AUTHORS Guillot,P., Schledzewski,K., Gratchev,A. and Goerdts,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-MAR-2002) Dermatology, Universitätsklinikum Mannheim,  
 Universitat Heidelberg, Theodor Kutzer Ufer 1-3, Mannheim, BW  
 68135, Germany  
 REMARK sequence update by submitter  
 COMMENT On Mar 25, 2002 this sequence version replaced gi:19568177.  
 FEATURES Location/Qualifiers  
 source 1..8157  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /tissue\_type="liver"  
 62..7741  
 /note="endothelial protein"  
 /codon\_start=1  
 /product="stabilin-2"  
 /protein\_id="AA191684.2"  
 /db\_xref="GI:19705588"  
 /translation="MARSKLLGLLPLILFLGLLVQNASPTPEAFELTKRCKKST"

LTIKTCQSSVNTAVKCPDGYIKITNGTVGRDVRSLKLSQSVLDIPGRHICRKO  
 YLQPCCFPHWPCMECPGARAPCGGRGVCDEMEGTGKSCSRAGRGFTACNCA  
 DVFPCNCSAYSCVHGNSG1S1GDTGCELSAYRGRCDKPIPECAALCPENSR  
 SPSSKDETKLCKLPSYKGGQCTKPIPCNKVCHPHASCSVLGNRHSVCQKGY  
 QGDQVCLPCTSPGCTSTKSVRYDGGCHCEKHEYNFVPGVSGMWDIC  
 ESKNPKHANCSTVSPGCTCTCKGVGGLNCSYGNIMORLRELNETPRGMWGOL  
 TSFISILDRITYAPLNSLPTVLLPSDKGLKGVYDVELLMKSAARFVYKLIAGQ  
 MSTQMYNLDITUTGSGEILINKDKQDKLKLKLYGSKIYQIIOGNIIVASNGLVHL  
 DRAMDKIEPTLESNQITIMTLOPRYKFRSLLEKTNVQALKEGKDEPYTIFVPS  
 NEALSNTAGVLDVLLSPGSKLLELYHYIYVAPTOLEVALTVLPHRCNETRKMILIT  
 FNISSKGOILANNVADTEVAANKNGRIYTLTGLVILPSPILPLPHRCNETRKMILIT  
 TCVRFCMKNPCSPNSPTALFTNCKFYGSRAWNLKIGCARVCDVVEI PRCKGFF  
 GPCNCPKGMPCNSGNGQCTIDELGNGTCTCEDGFGSFCQPCSKENRYGPOCNET  
 CQVHGICDRLSDGSLCTCRBTAGTRFCDKQTSACGPFMGFECHIHATCESNET  
 ASCVNDGEGDGLSCDKKPCGTSRGGSPNABCKOASTGTVCVQCRQGTNGR  
 DCVINSCLLPSGGCHNATCLYVGPQNECECKGPRNGIDCEPILEQCTBKC  
 HPLATQVTLTSGVSCVQEGVNGVLCYGNVLMELSFLEAAVYFQWIMNASIQSM  
 LSATSNLTVLPSLQAIKMDQNEKSFWLSRNNTPALIKYHTLITGYVADLQTLPS  
 HMLATSLGSEFLRLDKADNITIEGASVVDGDNATNGVHILINKVILPQRLTGLSLP  
 SILTLREMPDYSIFRGYIHYNLASALEADAYTVFPNNEALIESYIREKKATSLKE  
 DILQHVVLGKLLDHLNGHRETMIGFSYLLAFHLNDOLYNEAPINTVNTAT  
 DILQHVVLGKLLDHLNGHRETMIGFSYLLAFHLNDOLYNEAPINTVNTAT  
 KGVHGLSKVLEIKKRNNDNTIIVRGCKSQOCTLCPLETSLSETRKCIYSVYF  
 MGRSIFITGCOLQCVRTITISACCAFPQPCQACPCGKQVCSNGFCLDGVNGTGT  
 CECEGNGTACETCEGKGIHCDQASCVHRCNOCPSGDCDCDVGWKGKCDK  
 EITDNCNGTCHTSANCLLDPGKASCKAAGFGNGTCTAINACETISNGSCSAKAD  
 CKRTIPRSRVCVKAGYGDGIVCLINPCLNENHGCDRHAECTGTGPNQVNCPLPK  
 YTGDKVCTLLNVLCTNNGGSPFAPCNHTQDQRTCTCKPDYTGDDGIVCRGSHSEL  
 PKNPSTSYFQLOQSHAVOELAGPFTVPFSSDSFNSKLVKWDQGLMSQLIRY  
 HVACQQLLENLKVITISATLQGEPIISVANSQDTLVINKKAKVLSIIITNGVIHV  
 DITLSPONLLITPKASGRVLLNTTVAANHYTKFSKLIQDSLLKVIITDPMTVP  
 TLPWPTKALQALPOQDPLFNEQDKKAYLKFHVIRDTMALASDLPRSAWKTLL  
 QGSLSVRCGSDVGEFLFNGQCRILQRLIFDGGVAYGIDCLLMDTPGRCRDTF  
 TTFNIPGCGSGDVEFLPLOSCKVGRKCIYNLPFRDVEGVCNCTIIVHVPRC  
 CSYPMPCQACQGPDTPCNNRNGCYDQYKPTQCQCHTFNGFACSLCLPGRGPD  
 CQPCSHGQDEGITSQGLCEAGWTGRFCDAPTIVIPVCI PACSMHATCMENNT  
 CVCLNLYGDDGICTVVPFCQNGGCAKVAKCSQKQVSCQKQYKGDGHSQTEI  
 DPCANLYGDDGICTVVPFCQNGGCAKVAKCSQKQVSCQKQYKGDGHSQTEI  
 DPANCVLHQDTTGVVPHLSPLQYKLTIDKAKCAKAASTIATYNLSYAKAK  
 YHLCSAGLESGRAVPTIYASKCANITGVYDTRTNKSEMNDVFCRMRKDVNCTC  
 KAGVYDGFSCNGLLOVLMFSPSLNFTLVEVLSRASAQRAFLKHLTDLISGTL  
 FVQNSGLPKNKSLEGRIDIEHLITNVNVSFYDLDVLTGLKLSQLLITSSQDLH  
 QEARVDCRALOWDIIASNGVHLIIEPLKAPPTAATAHSGLTGTGFCVAVLVTGA  
 IALAAYSVERLNOTRTGPRPESEDDI DALAFKQKQPSINPLNYETSTPAAPBPSCD  
 PFTDSGERELNSDFLGALS"

ORIGIN

Alignment Scores:

Pred. No.: 6,89e-136 Length: 8157  
 Score: 1487.00 Matches: 275  
 Percent Similarity: 86.97% Conservative: 32  
 Best Local Similarity: 77.90% Mismatches: 44  
 Query Match: 79.73% Indels: 2  
 DB: 10 Gaps: 2

US-09-466-778B-11 (1-353) x AF364951 (1-8157)

QY 1 MetThrGlyProGlyHisHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20  
 DB 6512 ATGACGGGGCCCAAGCAAGTGTGAATGTAAGTAAAGTCACTATGTCCGGGATGGAAG 6571  
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40  
 DB 6572 GACTGTGAGCCCAACAGCTGCCCCCTTGACCGCTGCTTACAGACCAACGCAATGCCAC 6631  
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60  
 DB 6632 CCAGATGCCAACTGTGTGACCTTCATTCCAGACACTACTGTTCGGCGTATTCACCTTA 6691  
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGlnAlaCysAlaAsn 80  
 DB 6692 CGCTCCCCACTGGGCCAGTACAACTGACATTTGACAAAACCAAGCCCTGTGTGTAAG 6751  
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr\*\*GlnLysAlaLysTyrHis 100

DB 6752 GAAGCTGATCCATAGCACCTCAACACAGCTCTCCCTATGCCCAAGAGCCCAAGTATCAC 6811  
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120  
 DB 6812 CTCTGCTCAGCTGCTGGTGGAGAGTGGGGGGTTCCTTACCCACGATCTACGCTCTCT 6871  
 QY 121 GlnAsnCysGlySerGlyValValGlyValLeuValAspTyrGlyProArgProAsnLysSer 140  
 DB 6872 AAGAAATGT--GCAAAACATTTGCTGGGATTTGTAGACTACGGAACACGAGCAACAAGAGT 6928  
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr\*\*LysValGly 160  
 DB 6929 GAATGTGGAGAGCTTCTTGTTCAGATGAAAGATTTAATGCACCTGCAGCAGCAGC 6988  
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180  
 DB 6989 TACGTGGCGGATGGCTTCTCCTGCAATGGGAACCTGCTGCAAGTCTCATGCTCTCTCC 7048  
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200  
 DB 7049 TCATCTCAGAACTTCTCTGACAGAGGTGCTGCTTTTTCAGGAGCTCAGCCCAAGGCCGG 7108  
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220  
 DB 7109 CGCTTTTGAACACCTTGACTGCTGCTCATCAGTGGCACCTGTTTGTGCCACAGAAC 7168  
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240  
 DB 7169 AGTGGGCTACCGAAAAATAAGAGCTGTCTGGGGGGACATTGAGCACCACTCACTAAT 7228  
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260  
 DB 7229 GTCAACGCTCTCTTTTACGATGACCTGTCAATGATGACCTGCTGAGGAGCTAGGCTGGGA 7288  
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280  
 DB 7289 AGCCAACTGCTCATCACCTCCAGCAGCAGCAGCAGCTCCAC---CAAGAGCCAGGTTCTGTG 7345  
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValle 300  
 DB 7346 GATGGAAGAGCCATTTCTGACGTGGGACATCATTTGCTCTAAAGGGGTTCTCCATATCAT 7405  
 QY 301 SerArg\*\*LeuLysAlaProAlaProValThrLeu\*\*HisThrGlyLeuGly\*\*\* 320  
 DB 7406 TCTGAACCTTTGAAGAGCTCTCCAGCGCCGCAAGCTGCCACTCTGGCCCTAGGAACA 7465  
 QY 321 GlyIlePhe\*\*\*\*IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340  
 DB 7466 GGCATATTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7525  
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe\*\*\*HisPhe 353  
 DB 7526 TTCCGGCTAAACCGAGAACCAACTGGCTTCCGGCGGTTT 7564

RESULT 15

AX149447  
 LOCUS AX149447  
 DEFINITION Sequence 3 from Patent WO0136638.  
 ACCESSION AX149447  
 VERSION AX149447.1 GI:14347929  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Shinkets, R.A., Lichenstein, H., Vernet, C. and Fernandes, E.  
 TITLE Polypeptides and nucleic acids encoding same  
 JOURNAL Patent: WO 0136638-A 3 25-MAY-2001;  
 FEATURES  
 Location/Qualifiers  
 source 1..2011

Job time : 4966 secs

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores: 1.16e-108 Length: 2011  
Pred. No.: 1203.00 Matches: 230  
Score: 1203.00  
Percent Similarity: 84.87% Conservative: 0  
Best Local Similarity: 84.87% Mismatches: 4  
Query Match: 64.50% Indels: 37  
DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x AX149447 (1-2011)

```

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 1309 ATGACAGCCCGCGGCAGCAAGTGTGAGTGAAGTCACTATGTCGGAGATGGGCTG 1368
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 1369 AACTGTGAGCGGAGCAGCTGCCATTGACCGCTGCTTACAGGACAATGGCGAGTGCAT 1428
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 1429 GCAGAGCCAAATGTGGCGACTCCACTCCAGGATACCACTGTGGGTGTTCCATCTA 1488
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 1489 CGCTCCCGCAGTGGCCAGTATAAGCTGACCTTTCACAAAGCCAGAGAGGCGCTGTGCCAAC 1548
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 1549 GAAGCTGGACCATGGCAACCTACCAACAGCTCTCTATGCCCAAGGCCAAGTACCCAC 1608
Qy 101 LeuCysSerAlaGlyTyrLeuGlnThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 1609 CTGTGCTCAGCAGGCTGGCTGGAGACGGGGGTGGCTACCCACAGCCCTTCGCCCTCC 1668
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 1669 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGAGCTATGGACCCAGACCCCAACAGAGT 1728
Qy 141 GluMetTyrAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValG1 160
Db 1729 GAAATGTGGGATGTCTTCTGCTATCGGATGAAG----- 1762
Qy 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr 180
Db 1762 ----- 1762
Qy 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
Db 1763 -----GAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGGCG 1800
Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
Db 1801 TGCATTTCTAGAACACCTCGACTGACCTGTCCATCGCGGACCCCTCTTTGTGCCACAGAA 1860
Qy 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
Db 1861 CAGTGGGCTGGGGGAGATGAGACCTGTCTGGGGGGGACATCGNACACCACTCGCCNA 1920
Qy 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1 260
Db 1921 TGTACAGCATGTTTTCTACAATGACCTTGTCAATGGCACCACTGCAACCGAGGCTGG 1980
Qy 260 ySerLysLeuLeuIleThrAspArgGlnAsp 270
Db 1981 AAGCAAGCTGTCTCATCTGCCAGCCAGGAC 2011

```

Search completed: April 1, 2004, 11:22:00